

UNCLASSIFIED

NOV 77 A M OSTROWSKI

DA-ERO-75-G-035
NL

| OF |

AD
A050660

[illegible]

END
DATE
FILMED
4 -78
DDC

4 -78

AD A 050660

12

TOPICS IN THE THEORIE OF
ALGEBRAIC EQUATIONS

FINAL TECHNICAL REPORT

Alexander M. Ostrowski

November 1974 - November 1977

EUROPEAN RESEARCH OFFICE

United States Army

Contract Number DA-ERO-75-G-035

jointly with
Swiss National Science Foundation
Berne

Mathematical Institute, University of Basel
Switzerland

DDC
RECEIVED
MAR 2 1978
D

Approved for public release; distribution unlimited

AD No. 1
DDC FILE COPY

UNCLASSIFIED

R&D 2125

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
6	9	
4. TITLE (and Subtitle)		5. TYPE OF REPORT & PERIOD COVERED
TOPICS IN THE THEORIE OF ALGEBRAIC EQUATIONS •		FINAL TECHNICAL REPORT, NOV 74 - NOV 77
6. AUTHOR(s)		7. PERFORMING ORG. REPORT NUMBER
Alexander M. Ostrowski		
8. PERFORMING ORGANIZATION NAME AND ADDRESS		9. PROGRAM ELEMENT, PROJECT, AREA & WORK UNIT NUMBERS
Mathematical Institute University of Basel Switzerland		6.11.02A-1T61102B17C-00-497 1T161102B17C
10. CONTROLLING OFFICE NAME AND ADDRESS		11. REPORT DATE
U.S. Army R&S Group (EUR) Box 65 FPO NEW YORK 09510		NOV 77
12. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		13. NUMBER OF PAGES
		17 169 p.
14. DISTRIBUTION STATEMENT (of this Report)		15. SECURITY CLASS. (of this report)
APPROVED FOR PUBLIC RELEASE DISTRIBUTION UNLIMITED		UNCLASSIFIED
16. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		18a. DECLASSIFICATION/DOWNGRADING SCHEDULE
17. SUPPLEMENTARY NOTES		
18. KEY WORDS (Continue on reverse side if necessary and identify by block number)		
Algebraic Equations; Elimination; Kronecker-Lasker; Subdominant roots; Newton-Sylvester sign rule; Matrix Algebra; Numerical Methods.		
19. ABSTRACT (Continue on reverse side if necessary and identify by block number)		
SEE OVER		

FORM

1 JAN 73 1473

EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

20 Abstract

Contents:

- a. Applicability of the Bernoulli-L'Hospital rule was investigated, if the usual assumptions fail on a zero set.
- b. The probably most general extension of integrals of Cauchy-Frullani type was developed, which contains, beyond all special cases as yet known, a great number of further types of integrals.
- c. The classical theorem on irreducibility of the resultant of two general polynomials is extended to the case where one of these polynomials contains only one free parameter.
- d. An investigation was made of the properties of the Kronecker extensions of polynomial ideals.
- e. In discussions of the de Moivre - Laplace formula in the calculus of probabilities, ~~an error term is used. We prove that it is false and determine and investigate the correct term.~~ *and it is proved that the error term*
- f. Distributions of irrational numbers in a linear interval are investigated, ^{with} The results ~~are~~ extended to vectors with irrational coefficients in an m -dimensional interval in R^m .

R to the mth power

** is determined and investigated*

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

TOPICS IN THE THEORIE OF
ALGEBRAIC EQUATIONS

FINAL TECHNICAL REPORT

Alexander M. Ostrowski

November 1974 - November 1977

EUROPEAN RESEARCH OFFICE

United States Army

Contract Number DA-ERO-75-G-035

jointly with
Swiss National Science Foundation
Berne

Mathematical Institute, University of Basel
Switzerland

Approved for public release; distribution unlimited

ADDITIONAL TO	
DTIC	White Section <input checked="" type="checkbox"/>
DDC	Grey Section <input type="checkbox"/>
UNANNOUNCED	<input type="checkbox"/>
JUSTIFICATION	
BY	
DISTRIBUTION/AVAILABILITY CODES	
DISC.	AVAIL. and/or SPECIAL
A	

DDC
RECEIVED
MAR 2 1978
D

Table of contents

page

I.	Bernoulli-L'Hospital rule .	4
II.	Cauchy-Frullani integrals .	5
III.	Irreducibility of the resultant and connected results .	6
IV.	Algebraic closure of modules.	7
V.	On Kronecker's elimination theory.	8
VI.	On the remainder term of the de Moivre-Laplace formula.	10
VII.	Distribution function of certain sequences mod 1 .	12
VIII.	Rational approximations to an irrational number .	12
IX .	The Error term in multidimensional diophantine approximation.	13

Abstract

- I. Applicability of the Bernoulli-L'Hospital rule was investigated, if the usual assumptions fail on a zero set.
- II. The probably most general extension of integrals of Cauchy-Frullani type was developped, which contains, beyond all special cases as yet known, a great number of further types of integrals.
- III. The classical theorem on irreducibility of the resultant of two general polynomials is extended to the case where one of these polynomials contains only one free parameter.
- IV. While a polynomial ideal, P , with a basis P_ν ($\nu=1, \dots, n$) is defined as the set of all polynomials of the form

$$\sum_{\nu=1}^n K_\nu P_\nu, \quad K_\nu \text{ polynomials,}$$

the Kronecker extension of the ideal P , \bar{P} , is defined as the set of all polynomials satisfying an equation of the type

$$\pi^m + \sum_{\mu=1}^m K_\mu \pi^{m-\mu} = 0$$

where each K_μ is a polynomial from P^μ . The investigation was concerned with the properties of the Kronecker extensions of polynomial ideals, defined in this way.

polynomial ideal under consideration. This invariance can be partly saved introducing algebraically closed ideals as discussed in the technical report BMN 46.

VI. In different discussions of the de Moivre - Laplace formula in the calculus of probabilities, from Laplace, 1812, to Feller, 1950, an error term was used of the form $\frac{d_n}{\sqrt{2\pi n}} e^{-\frac{1}{2}} + O(\frac{1}{n})$ and it was asserted, that d_n can be taken as 1. We prove that this is false in the sense, that such a form of the remainder is possible, but the d_n , instead of being = 1, are, with $n \rightarrow \infty$, everywhere dense between -1 and 1.

VII. The investigation is concerned with the problem whether the expressions d_v used in the abstract VI are uniformly distributed in $(-1, 1)$. We prove that this is not the case and obtain explicit expressions for the density of d_v and some more general sequences.

VIII. If J is a linear interval mod 1 of the length $|J|$, α a real irrational and $R(v\alpha)$, $v = 1, 2, \dots$ are the residuals mod 1 of the products $v\alpha$, then $N(n, J)$, the number of the $R(v\alpha)$ from J with $v \leq n$, satisfies the relation

$$(1) \quad N(n, J) = n|J| + E(n), \quad E(n) = o(n) \quad (n \rightarrow \infty).$$

The article investigates the improved estimates of $E(n)$ for special α , in particular in connection with $T(n)$ defined by

$$(2) \quad T(n) := \min |z_1 \alpha + z_0| \quad (|z_1| \leq n, \quad z_0 \wedge z_1 \in \mathbb{Z}).$$

The central result is a functional inequality implying the functions $E(n)$ and $T(n)$.

IX. The aim of the investigation is to generalize the results mentioned in Abstract VIII to the case of \mathbb{R}^m , $m > 1$, where the irrational α is replaced by a vector $\alpha := (\alpha_1, \dots, \alpha_m)$ satisfying the corresponding independency condition.

lying mod 1 in J with $1 \leq v \leq n$. The Error function $E(n)$ is again defined by

$$(1) \quad E(n) := N(n, J) - n|J| = o(n) \quad .$$

On the other hand, the functions $\Psi(\sigma)$, $\varphi(\tau)$ and $A^*(x)$ are defined similarly as in BMN 50, while the corresponding functional inequality has to be written as

$$(2) \quad A^*(x) \leq \alpha A^*(\tau) + Bx / \varphi\left(\frac{1}{\gamma_0 \tau}\right), \alpha > 1, B > 0,$$

for a convenient constant γ_0 . Different solutions of the functional inequality (2) are discussed and corresponding estimates of $E(n)$ obtained.

I. Bernoulli-L'Hospital rule

In the usual formulations of the general Bernoulli-L'Hospital rule one of the four limiting processes

$$(1) \quad x \uparrow \infty, \quad x \downarrow -\infty, \quad x \uparrow x_0, \quad x \downarrow x_0$$

is considered. The functions $f(x)$, $g(x)$ are assumed to have derivatives, where g' is either always > 0 or always < 0 , while $|g| \rightarrow \infty$. Then the assertion is (see the appended BMN 43)

$$(2) \quad \liminf f'/g' \leq \overline{\lim} f/g \leq \overline{\lim} f'/g'.$$

It is shown on a counter example that this formulation is no longer true, if the existence of f' and g' fails for a zero set, Ω , while such an exceptional set often occurs if for instance f and g are Lebesgue integrals.

It can then be shown that, if both f and g are absolutely continuous, but f' and g' only exist with the exception of a zero set, Ω , - and again g' has a fixed sign (save on Ω) and $|g| \rightarrow \infty$ - then (2) again holds if Ω is disregarded in the extreme terms of (2).

As a matter of fact, we obtain still a partial result if we allow g to be discontinuous. Then we have at least

$$(3) \quad \liminf \min(0, f'/g') \leq \overline{\lim} f/g \leq \overline{\lim} \max(0, f'/g').$$

A particularly useful rule is obtained in the following result :

Assume that for one of the limiting processes (1), save on a set, Ω , of measure 0 in the range of x , $f'(x)$ and $g'(x)$ exist, $g'(x)$ is either always > 0 or always < 0 and $g(x)$ tends monotonically to $+\infty$ or $-\infty$. Assume further that we have for a finite constant α ,

$$(4) \quad f'(x)/g'(x) \longrightarrow \alpha, \quad |\alpha| < \infty \quad (x \notin \Omega),$$

and that $f(x) - \alpha g(x)$ is absolutely continuous. Then the relation holds:

$$(5) \quad f(x)/g(x) \longrightarrow \alpha.$$

All details are contained in BMN 43 distributed previously. The paper appeared in the American Mathematical Monthly, 83, 1976, pp. 239-242.

II. Cauchy-Frullani integrals

The Cauchy-Frullani formula is

$$(1) \quad \int_0^{\infty} [f(at) - f(bt)] dt/t = [f(\infty) - f(0)] \lg a/b \quad (a \wedge b > 0).$$

This holds if f is L integrable and the two expressions on the right, defined by

$$(2) \quad f(\infty) := \lim_{t \uparrow \infty} f(t), \quad f(0) := \lim_{t \downarrow 0} f(t)$$

exist.

It is natural to ask whether $f(\infty)$ and $f(0)$ can be replaced, if these limits do not exist, by convenient mean values, $M(f)$ and $m(f)$. Indeed, it can be shown, that if both limits

$$(2^0) \quad M(f) := \lim_{x \uparrow \infty} \frac{1}{x} \int_1^x f(t) dt, \quad m(f) := \lim_{x \downarrow 0} x \int_x^1 f(t) dt/t^2$$

exist, then the formula

$$(3) \quad \int_0^{\infty} [f(at) - f(bt)] dt/t = [M(f) - m(f)] \lg a/b$$

holds, and vice versa, if the integral in (3) converges at least for a set of positive measure of the quotients a/b , then the expressions (2^0) exist and (3) holds for all positive a and b . (See the chapters I - III and V of *BM* 44.)

As a matter of fact, necessary and sufficient conditions of this kind were indicated in 1940 by K.S.K. Jyengar, the conditions corresponding to $f(\infty)$ being

$$(4) \quad \exists \int_1^{\infty} f(t) dt/t^2, \quad \exists \lim_{x \rightarrow \infty} x \int_x^{\infty} f(t) dt/t^2,$$

where the second limit corresponds in (1) to $f(\infty)$. Jyengar's conditions corresponding to $f(0)$ are similar.

However, Jyengar's proof was incorrect. A simpler proof was given 1942 by Agnew, but this proof contains again a gap, as was pointed out 1949 by A. Ostrowski who gave the conditions (2^0) and sketched a direct proof

of their equivalence with Jyengar' conditions. This proof is detailed in the chapter IV of BMN 44. Five years later, 1954, Agnew succeeded in filling out the gap of his proof.

The following sections of the paper contain the formulation and the proof of a very general extension of (1) and (3), the Three-Functions-Formula. Its formulation contains, because of its generality, several parameters, and it is sufficient to refer to the section 45, pp.36-37, of the BMN 44. To characterize the formula we give in the following a simplified version of it.

Assume $G(x)$ L -integrable in $(0, \infty)$, bounded in any closed interval of the positive x -axis and such that $M(G)$, $m(G)$ exist. Assume the open interval J between a and b ($a \leq b$) and two functions $\varphi(t)$, $\psi(t)$ defined and absolutely continuous in J . Assume further that the values $\varphi(a)$, $\psi(a)$, $\varphi(b)$, $\psi(b)$ exist if defined as limits from J , and further that $\varphi'(a)$, $\psi'(a)$, $\varphi'(b)$, $\psi'(b)$ exist and $\psi'(a)\psi'(b) \neq 0$.

Then we have the formula

$$(5) \int_a^b \{ \psi'g(\psi)/\psi - \varphi'g(\varphi)/\varphi \} dt = m(g) \varphi \frac{\psi'(a)}{\psi'(a)} - M(g) \varphi \frac{\psi'(b)}{\psi'(b)}.$$

All details are contained in BMN 44, distributed previously. The paper has appeared in *Commentarii Mathematici Helvetici*, 51, 1976, pp.57-91.

III. Irreducibility of the resultant and connected results

Consider two polynomials $f(x)$ and $g(x)$ with coefficients respectively a_μ and b_ν , and their resultant R . Assume that the b_ν are polynomials in a parameter ξ so that $g(x)$ becomes a polynomial $g(x, \xi)$. Then, if $g(x, \xi)$ is irreducible, R as a polynomial in the a_μ and ξ is also irreducible.

As corollaries of this result, equations satisfied either by the sums of zeros of f and g or by the products of these zeros are analyzed and their irreducibility is proved.

All details of the proofs are contained in BMN 45, distributed previously. The paper is in print in *Archiv der Mathematik*.

IV. Algebraic Closure of Modules

While a polynomial ideal, P , with a basis P_ν ($\nu=1, \dots, n$) is defined as the set of all polynomials of the form

$$\sum_{\nu=1}^n K_\nu P_\nu, \quad K_\nu \text{ polynomials,}$$

the Kronecker extension of the ideal P , \bar{P} , is defined as the set of all polynomials satisfying an equation of the type

$$\pi^m + \sum_{\mu=0}^m K_\mu \pi^{m-\mu} = 0$$

where each K_μ is a polynomial from P^μ .

The technical report BMN 46, the copies of which were appended to the periodic technical report of April 76, is concerned with the properties of the Kronecker extensions of polynomial ideals, defined in this way, and the generalization of this concept for the modules over a ring.

The main points of this discussion are

a) "Linearization" of the algebraic condition imposed on the elements, π , of \bar{P} . This is done, introducing the concept of a supporting sequence of π over P in theorem 1 in sec. 12 (partly due to Prüfer), and more generally, introducing the concept of a double supporting sequence in theorem 1⁰ in sec. 15.

We further generalize this criterion in theorem 7 in sec. 51, though useful, but bearing no longer linear character.

b) The proofs that \bar{P} is a module, theorem 2, sec. 17, and that \bar{P} is algebraically closed, $\bar{\bar{P}} = \bar{P}$, in theorem 4 of sec. 33.

c) An importante theorem of Macaulay, asserting that, if A, B, C are modules and B finite, then from $BC \subset \bar{AB}$ follows $C \subset \bar{A}$, is generalized,

in theorem 8, sec. 60, to modules over a ring, together with a considerable simplification of the proof.

A great number of partial results are more of technical character but useful and partly even necessary in applications of the theory. In particular, attention may be drawn to the discussion in the sec. 61-78, of the application of the valuation theory to rings and modules, which is very technical but proved to be important in the applications of the whole theory to Kronecker's theory of elimination, discussed in the next item. The paper is being printed in Crelle's Journal der Mathematik

V. On Kronecker's elimination theory

Kronecker's set up in his elimination theory starts with a polynomial ideal

$$(1) \quad P(f_\nu(x_1, \dots, x_n)) \quad (\nu = 1, \dots, n) \quad \text{in } \Omega[x_1, \dots, x_n]$$

where $\Omega := \mathbb{C}$ and the n polynomials f_ν with coefficients from Ω form a basis of P . To obtain the null manifold of P Kronecker makes first the variables x_ν to undergo a general linear transformation such that the highest powers of each single variable x_1, \dots, x_n in any of the f_ν has a non vanishing coefficient from \mathbb{C} . Then he eliminates first x_1 computing the resultant with respect to x_1 :

$$(2) \quad \text{Res}_{x_1} \left(\sum_{\nu=1}^n u_\nu f_\nu, \sum_{\nu=1}^n v_\nu f_\nu \right) = \sum_{\alpha, \beta} U_\alpha V_\beta R_{\alpha, \beta}.$$

Here the u_ν, v_ν are indeterminates independent with respect to

$\Omega[x_1, \dots, x_k]$. On the right side of (2) the U_α are products of powers of the u_α , the V_β products of powers of the v_β and $R_{\alpha, \beta}$ are independent of the u_α , v_β , while the products $U_\alpha V_\beta$ are all distinct. Then the module

$$(3) J_f := (R_{\alpha, \beta}(x_2, \dots, x_k)) \text{ in } \Omega(x_2, \dots, x_k)$$

is the resultant module obtained from P by elimination of x_1 .

However, J_f as defined starting from (2) is not invariant with respect to the choice of the basis elements f_α of P .

The first main result of our discussion in case $\Omega := \mathbb{C}$, is that, the ideal \bar{J}_f is indeed independent of the choice of the basis of P and as the matter of fact only depends on \bar{P} .

However, the whole discussion is undertaken on a larger front in so far as Ω need not be $= \mathbb{C}$ but is assumed only a natural ring. This embraces also the set up of Lasker in the case $\Omega = \mathbb{Z}$.

Our second main result can be formulated as follows. Let

$$(4) \quad Q = (g_\mu(x_1, \dots, x_k)) \quad (\mu = 1, \dots, m)$$

be another ideal in Ω such that $\bar{Q} = \bar{P}$ in Ω . Denote by α_f the ideal in $\Omega[x_2, \dots, x_k]$ the basis of which is formed by the coefficients of the highest power of x_1 in the n polynomials f_α , and α_g the corresponding ideal formed for the g_μ . Then, if we denote with J_g

the ideal in $\Omega[x_2, \dots, x_k]$ defined applying (2) and (3) to the e_μ , the relations hold

$$(5) \quad \overline{\alpha_g^{2r} J_g} \subset \overline{J_f}, \overline{\alpha_f^{2r'} J_f} \subset \overline{J_g}.$$

Here, r is any integer $\geq p$ and r' any integer $\geq q$, where p, q are correspondingly the maximal degrees of the f and of the g_μ in x_1 .

It is of some interest to obtain, if possible, lower values for r and r' . This is indeed possible, using the concept of the height of a polynomial ideal with respect to a Dedekind module as developed in the last part of the technical report BMN 46. However, this part of the discussion requires a not inconsiderable display of technical argumentations.

The technical report BMN 47 has been distributed previously. The paper is being printed in Crelle's Journal der Mathematik.

VI. On the remainder term of the de Moivre-Laplace formula

The de Moivre-Laplace formula* in the probability calculus is concerned with the case of the n times repeated trials with constant probability p , if $n \rightarrow \infty$.

*A de Moivre, Miscellanea analytica (2nd supplement), 1733;
P. Laplace, Oeuvres, Vol. VII, 1812, pp 281, 284.

Introduce , for a p , $0 < p < 1$, and integer v with $0 < v < n$, the notations :

$$p_v := \binom{n}{v} p^v (1-p)^{n-v}, \quad w := \sqrt{2p(1-p)n},$$

$$n_0 := \lfloor (n+1)p \rfloor, \quad w' := \sqrt{\frac{2n_0(n-n_0)}{n}}.$$

Then, for a positive η , the formula in question in the form in which it was given by Laplace can be written as

$$L(\eta) := \sum_v p_v (|v - n_0| \leq \eta w'),$$

$$L(\eta) = \frac{2}{\sqrt{\pi}} \int_0^\eta e^{-\tau^2} d\tau + \frac{d'_n}{\sqrt{\pi} w'} e^{-\tau^2} + O\left(\frac{1}{n}\right), \quad d'_n = 1,$$

while usually the corresponding formula can be written as

$$M(\eta) = \sum_v p_v (|v - np| \leq \eta w),$$

$$M(\eta) = \frac{2}{\sqrt{\pi}} \int_0^\eta e^{-\tau^2} d\tau + \frac{d_n}{\sqrt{\pi} w} e^{-\tau^2} + O\left(\frac{1}{n}\right), \quad d_n = 1.$$

The technical report BMN 48 contains the proof that both formulas are false as to the values of d'_n , d_n , and that even, if $n \rightarrow \infty$, both d'_n , d_n are everywhere dense between -1 and 1 .

Both d'_n and d_n can be expressed explicitly using the function $R(x) := x - [x]$. Such a representation was already given for d_n by Uspensky, 1937,* who however expressed himself very cautiously as to the assertion of Laplace.

The proof, that d'_n and d_n are everywhere dense in $(-1, 1)$ requires argumentation belonging to the theory of Diophantine approximations. The proofs are found in the technical report BMN 48 distributed previously.

* J. Uspensky, Introduction to mathematical probability,
New York, 1937.

VII. Distribution Function of certain sequences mod 1.

While the sequence

$$d_v := R(\eta\sqrt{2p(1-p)v} + pv) + R(\eta\sqrt{2p(1-p)v} - pv)$$

is, as proved under VI, everywhere dense in $(-1, 1)$, the present article discusses the problem whether it is uniformly distributed. We prove that this is not the case. The discussion is carried out for the general sequence

$$R(a\sqrt{v} + v\lambda) + R(a\sqrt{v} - v\lambda) \quad (v = 1, 2, \dots)$$

and even more generally the sequence

$$(1) \quad d_v := R(\alpha_v + v\lambda) + R(\alpha_v - v\lambda) \quad (v = 1, 2, \dots).$$

The sequence α_v in (1) is assumed to satisfy certain rather special conditions which are for instance satisfied for $\alpha_v = av^\alpha$, $0 < \alpha < 1$.

We prove that, for any irrational λ the density of (1) is x in any point x with $0 < x < 1$, and $1-x$ in any point x with $1 < x < 2$. In the case of a rational λ (1) is uniformly distributed if and only if 2λ is integer. We prove that the distribution function always exists for rational λ , too, and obtain this function in a neighborhood of 1.

The proofs require a rather intricate discussion of some integrals in connection with diophantine approximations. The complete paper, as the technical report BMF 49, containing all details of the proofs, has already been distributed. The paper is accepted for publication in Acta Arithmetica.

VIII. Rational approximations to an irrational number.

In the following J means an interval mod 1 of the length $|J|$, while $R(v\alpha)$, $v = 1, 2, \dots$ signify the residual mod 1 of the product $v\alpha$, for a fixed real irrational α . If, for an $n \geq 1$, $N(n, \alpha)$ is the number of the $R(v\alpha)$ from J with $v \leq n$, then

$$(1) \quad E(n) := N(n, J) - n|J| = o(n) \quad (n \rightarrow \infty),$$

by a result due to Bohl-Sierpinski-Weyl.

The aim of the investigation is to improve (1) under special assumptions for α . In this connection α will be characterized by the function $T(n)$ defined by

$$(2) \quad T(n) := \min |z_1 \alpha + z_0| \quad (|z_1| \leq n, z_0 \wedge z_1 \in \mathbb{Z}).$$

We establish a functional inequality depending in a certain way on $E(n)$ and $T(n)$. To the purpose we use an arbitrary strictly decreasing and continuous function $\Psi(n)$ such that

$$(3) \quad 0 < \Psi(n) \leq T(n), \quad \Psi(n) \downarrow 0$$

and the inverse of Ψ , φ , so that

$$(4) \quad \tau = \Psi(\sigma), \quad \sigma = \varphi(\tau).$$

On the other hand we introduce a majorant of $E(n)$ by

$$(5) \quad A^*(x) = \sup_{1 \leq n \leq x} E(n) \quad (J \text{ arbitrary}).$$

Then the functional inequality in question is

$$(6) \quad A^*(x) \leq A^*(\tau) + 4x/\varphi\left(\frac{1}{2\tau}\right) \quad (x \geq 1, \tau \geq 2).$$

In particular, from (6) can be easily obtained, for an α with bounded partial denominators in the continued fraction,

$$(7) \quad E(n) = O(\lg n),$$

and, for algebraic α ,

$$(8) \quad E(n) = O(n^r),$$

for a convenient number r with $0 < r < 1$.

The detailed proofs are contained in the appended report BII 50.

IX. The Error term in multidimensional diophantine approximation.

In generalization of the situation dealt with in VIII consider the space \mathbb{R}^m ($m > 1$) and an m -dimensional vector $\alpha := (\alpha_1, \dots, \alpha_m)$ satisfying the condition that always $z_1 \alpha_1 + \dots + z_m \alpha_m + z_0 \neq 0$ for integer z_μ with $\sum_{\mu=0}^m |z_\mu| > 0$.

Under the symbol $R(\alpha_\mu)$ we understand the vector

$$(1) \quad R(\alpha_\mu) = (v\alpha_1, \dots, v\alpha_m) \pmod{1}$$

while J has the meaning of an m -dimensional interval mod 1, the volume of which is denoted by $|J|$. The expression $N(n, J)$ signifies the number of $R(\nu\alpha_\mu)$ from J with $1 \leq \nu \leq n$.

Then the Error term of uniform distribution of the $R(\nu\alpha_\mu)$ in J is given by

$$(2) \quad E(n) := N(n, J) - n|J| = o(n) \quad (n \rightarrow \infty),$$

uniformly for all J .

We are concerned with the problem to improve the estimate (2) under special assumptions about the α_μ . To this purpose we introduce the norm $|z|_\infty$ of the integer vector $z = (z_1, \dots, z_m)$ by

$$|z|_\infty := \max_{\mu} |z_\mu|$$

and denote generally for all real number a by $|a|$ the distance of a from the nearest integer.

The vector α is to be characterized by a strictly diminishing function $\tau = \tau(\sigma) \downarrow 0$ such that

$$(3) \quad \Psi(\sigma) \leq \min_{\mu=1, \dots, m} \sum_{\nu=1}^{\sigma} z_\nu \alpha_\mu \quad (1 \leq |z|_\infty \leq \sigma, \sigma \geq 1).$$

Denote the inverse of $\tau = \tau(\sigma)$ by $\sigma = \varphi(\tau)$ and put

$$\Lambda^*(x) := \sup E(n) \quad (1 \leq n \leq x, J \text{ arbitrary}).$$

If we finally put

$$y_0 := 2^{m+1} / ((m+1)!)^2$$

the central result of the paper is the inequality

$$(4) \quad \Lambda^*(x) \leq \alpha \Lambda(\tau) + \beta x / \varphi\left(\frac{1}{y_0 \tau}\right), \quad \alpha > 1, \beta > 0.$$

The relation (4) is independent of the dimension m . The most interesting case is obtained under the assumption

$$(5) \quad \varphi\left(\frac{1}{y_0 x}\right) \geq \left(\frac{x}{l(x)}\right)^\vartheta, \quad l(x) \uparrow, \quad 0 < \vartheta < 1.$$

We obtain then

$$(6) \quad \Lambda^*(x) = O(x / \varphi\left(\frac{1}{y_0 x}\right)) \quad (x \uparrow \infty).$$

The essential problem in applying (6) is then to find special $\ell(x)$ for which the inverse function, Ψ , of φ can be simply expressed or well approximated.

In the special case $\ell(x) = \text{const.} = c > 0$ the condition corresponding to (5) is

$$(7) \quad x^{1/q} \Psi(x) \gg c_1 > 0$$

and here we obtain

$$(8) \quad \Lambda^*(x) = O(x^{1-f}) .$$

Such relations hold always if $\alpha_1, \dots, \alpha_n$ are algebraic numbers.

If the estimate obtained for $\ell(x) = \text{const.}$ is too rough we obtain finer estimates introducing a function $k(x)$ strictly monotonically going to ∞ with $x \rightarrow \infty$ and satisfying the condition

$$(9) \quad xk'(x) = o(k(x)/\lg k(x)) \quad (x \rightarrow \infty) .$$

Then, taking $r = 1/q > 1$ and subjecting $\Psi(x)$ to the condition

$$(10) \quad \gamma_c(cx)^r k((cx)^r) \Psi(x) > 1 \quad (x \gg x_0)$$

for a convenient constant $c > 0$, we obtain

$$(11) \quad \Lambda(x) = O(x^{1-f} k(x)^s) .$$

The technical report BMN 51 appended to this report contains the detailed proofs of the results indicated above as well as a discussion of some solutions of the inequality (4) which present a certain interest although they cannot be applied to the problem on diophantine approximations which is the main subject of the paper.

Appendices:

- | | |
|--------|--|
| BMN 50 | On rational approximations to an irrational number. |
| BMN 51 | On the Error term in multidimensional diophantine approximation. |

List of technical reports issued during
the grant period.

- | | |
|--------|--|
| BMN 43 | Note on the Bernoulli - L'Hospital Rule. |
| BMN 44 | On Cauchy - Frullani Integrals. |
| BMN 45 | The irreducibility of the resultant and connected irreducibility theorems. |
| BMN 46 | Algebraic closure of modules. |
| BMN 47 | On Kronecker's Elimination Theory. |
| BMN 48 | On the remainder term of the de Moivre - Laplace formula. |
| BMN 49 | On the distribution function of certain sequences (mod 1). |
| BMN 50 | On rational approximations to an irrational number. |
| BMN 51 | On the error term in multidimensional diophantine approximation. |

August 1977

BMN 50

BASEL MATHEMATICAL NOTES

ON RATIONAL APPROXIMATIONS TO AN IRRATIONAL NUMBER

by A.M.Ostrowski

Sponsored in part by the Swiss
National Science Foundation.

Sponsored in part under the Grant
DA-ERO-75-G-035 of the European
Research Office, United States Army,
to the Institute of Mathematics,
University of Basel.

Errata BMN 50
=====

p.1	,	5.1.f.b.	replace "by" with "generally by"
p.1	,	10.1.f.b.	replace "Chebyshev" with "Tchebyshev"
p.2	,		add to footnote 1 : My attention was kindly drawn to Khintchine's paper by professor Cassels.
p.3	,		replace in formula (1.9), " $\frac{1}{\psi(1)}$ " with " $\frac{1}{2\psi(1)}$ " replace in formula (1.12) " $\sigma \geq 2$ " with " $\sigma \geq \frac{1}{2\psi(1)}$ "
p.6	,	3.1.f.a.	replace "are lying" with "lie"
p.6	,	6.1.f.a.	add: , by (2.1),
p.8	,	6.1.f.a.	replace "(2.11)" with "(3.1)"
p.8	,	3.1.f.b.	replace "and" with "for a certain r_k and"
p.10	,	9.1.f.b.	replace " in sec.1 ... expressions" with "expressions defined in sec.1 and sec.4"
p.11	,	4.1.f.a.	replace "and since" with ".Since"
p.12	,	5.1.f.b.	replace "end" with "upper end"

1

On rational approximations to an
irrational number.

§ 1. Introduction .

1. Denote by $R(a)$, $0 \leq R(a) < 1$, the fractional part of a .
Further, for a real irrational α , denote by $N(n, J)$ the number of
the elements of the sequence $R(\nu\alpha)$ ($\nu = 1, \dots, n$) lying in a subintervall,
 J , of $\langle 0, 1 \rangle$ modulo 1. α remains fixed throughout .

In the case of a real irrational number, α , the two following
results are classical.

a) The sequence $R(\nu\alpha)$, $\nu = 1, 2, \dots$ is everywhere dense in
the half open intervall $\langle 0, 1 \rangle$ (Chebyshev) .

b) The sequence $R(\nu\alpha)$, $\nu = 1, 2, \dots$ is uniformly distributed
in $\langle 0, 1 \rangle$ (Bohl [1] , Sierpinski [2] , Weyl [1]). This signifies that
for any subintervall, J , of $\langle 0, 1 \rangle$, the relation holds:

$$(1.1) \quad N(\nu, J) = \nu|J| + o(\nu) \quad (\nu \rightarrow \infty),$$

denoting by $|J|$ the length of J .

2. As to the result a) the question arises how far we must go
in the sequence of the $R(\nu\alpha)$ in order to obtain an approximation to ξ
with an error $\leq \epsilon$ for any ξ from $\langle 0, 1 \rangle$. More precisely, we seek to
define a function $\Phi(\epsilon)$, tending monotonically to ∞ with $\epsilon \downarrow 0$ such that

for any ξ from $(0,1)$ there exist two integers $x > 0$ and y , such that

$$(1.2) \quad |x\alpha - y - \xi| \leq \xi, \quad 1 \leq x \leq \bar{\Phi}(\xi) .$$

3. In connection with this problem we have the mention an important and extremely general theorem by Khintchin which, if specialized to our problem, gives a solution of a similar problem in which the condition $0 \leq x \leq \bar{\Phi}(\xi)$ is replaced by the condition

$$(1.3) \quad |x| \leq \bar{\Phi}(\xi) \quad ^1).$$

The function $\bar{\Phi}_0(\xi)$ is expressed in terms of the function $T(\sigma)$ defined by

$$(1.4) \quad T(\sigma) = \min |z_1\alpha + z_0|, \quad |z_1| \leq \sigma, \quad z_0 \wedge z_1 \in \mathbb{Z},$$

and an arbitrarily chosen continuous and strictly monotonically decreasing function $\Psi(\sigma)$ satisfying the relation

$$(1.5) \quad 0 < \Psi(\sigma) \leq T(\sigma) \quad (\sigma \geq 1) .$$

We obtain then

$$(1.6) \quad \bar{\Phi}_0(\xi) := \frac{1}{16\Psi(1/16\xi)} .$$

and this expression is "the best" save for the values of constants.

In this note we obtain, by a very elementary discussion, for $\bar{\Phi}(\xi)$ the expression

1) Khintchin [1], see also Cassels [1] pp. 97 - 99.

$$(1.7) \quad \bar{I}(\varepsilon) = \frac{1}{2\Psi(3/\varepsilon)} .$$

This expression is derived in sec.10. It is obviously "the best" save for the values of constants.

4. As to the problem b) our solution of this problem depends on the function $\Psi(\sigma)$ introduced in (1.5). Denote the inverse function of $\tau = \Psi(\sigma)$ by

$$(1.8) \quad \varphi(\tau) = \sigma, \quad (\tau \geq \Psi(1)) , \quad \varphi(\Psi(\sigma)) = \sigma \quad (\sigma \geq 1) .$$

This is allowed as $\Psi(\sigma)$ is continuous and strictly monotonically decreasing. Put further

$$(1.9) \quad U(\tau) := \frac{2}{\varphi\left(\frac{2}{\tau}\right)} \quad (\tau > \Psi(1))$$

and for any partial intervall J of $(0,1)$:

$$(1.10) \quad A(n,J) := n |J| - N(n,J) ,$$

$$(1.11) \quad A^*(x) := \sup_{\substack{1 \leq n \leq x \\ J}} A(n,J) .$$

Our central result is the following functional inequality for $A^*(x)$:

$$(1.12) \quad A^*(x) \leq A^*(\sigma) + 2xU(\sigma) \quad (x > 0, \sigma \geq 2) .$$

This formula is derived in § 5.

5. Using the relation (1.12) we can easily prove that for all α with bounded partial quotients of their continued fraction expansion, and in particular for all quadratic irrationals α the relation holds :

$$(1.13) \quad A^*(x) = O(\lg x) \quad (x \rightarrow \infty) .^2)$$

Another important case is that in which, for two constants g and g' :

²⁾ First proved in Hardy and Littlewood [2] and Ostrowski [1] .
See also Behnke [1] .

$$(1.14) \quad U(x) \leq g' U(gx) \quad , \quad \varepsilon > g' > 1 \quad (x \geq x_0 > 0).$$

4

In this case we obtain

$$(1.15) \quad A^*(x) = O(xU(x)) \quad (x \rightarrow \infty).$$

The condition (1.14) can be used immediately for instance for

$$(1.16) \quad \psi(\sigma) = \frac{c}{\sigma} \quad , \quad c > 0 \quad , \quad \sigma > 1 \quad .^3)$$

6. The method used in our discussion of the problem b) can be considered as a further development of a method used by E. Hecke, Hecke [1] pp. 331 - 335. Hecke made in particular essentially use of the expression

$$(1.17) \quad S(n, \xi) := \sum_{\nu=1}^n R(\xi + \nu\alpha) \quad ,$$

however, only in the special case $\xi = 0$. (I had occasion, during writing down of the above quoted paper, to add some remarks which were then incorporated, with due credit, in Hecke's paper, l.c. pp. 332, 335.)

However, Hecke did only arrive at partial results. For instance in the case of bounded partial quotients he says that he could only obtain the estimate $O(e^{\sqrt{21} \lg x})$ instead of $O(\lg x)$.

It may be finally mentioned that an inequality similar to and a little weaker than (1.12) can be deduced by another method (cf. Ostrowski [3]) for an essentially more general case for the n dimensional approximations. See our forthcoming communication: On n dimensional approximations.

³⁾ The corresponding result was first proved in Ostrowski [1]; see also Hecke [1].

5

§ 2. Derivation of (1.7) .

7. Observe that in (1.4),

$$(2.1) \quad \Psi(\sigma) \leq T(\sigma) \leq \frac{1}{2\sigma} ,$$

by an inequality going back to Dirichlet (see Cassels [4], theorem 1, p.1).

Applying the monotonically decreasing function φ it follows $\sigma \geq \varphi\left(\frac{1}{2\sigma}\right)$,

$$(2.2) \quad \frac{1}{\sigma} \leq \frac{1}{\varphi\left(\frac{1}{2\sigma}\right)} = U(\sigma) \quad \left(\sigma \geq \frac{1}{2\varphi(1)}\right) .$$

8. Choose now an arbitrary ξ from $\langle 0,1 \rangle$. From (1.4) it follows that for a certain positive integer $q \leq \sigma$,

$$T(\sigma) = |q\alpha - p| , \quad 1 \leq q \leq \sigma ,$$

with $(q,p) = 1$ and therefore

$$(2.3) \quad p = q\alpha \pm T(\sigma) .$$

Consider now the congruence

$$(2.4) \quad xp \equiv [q\xi] = q\xi - R(q\xi) \pmod{q} ,$$

which has as a solution a positive integer x between 1 and q . Eliminating in (2.4) p by (2.3), we obtain

$$q(x\alpha - \xi) \equiv \pm xT(\sigma) - R(q\xi) \pmod{q}$$

and dividing both sides by q ,

$$x\alpha - \xi \equiv \pm \frac{xT(\sigma)}{q} - \frac{R(q\xi)}{q} \pmod{1} .$$

It follows that, for a convenient integer p_1 ,

$$x\alpha - \beta - p_1 = \theta T(\sigma) + \frac{\theta}{q}$$

where $|\theta|$ and θ are lying between 0 and 1,

$$(2.5) \quad -T(\sigma) < x\alpha - \beta - p_1 \leq T(\sigma) + \frac{1}{q}.$$

9. In order to obtain an upper estimate of V_q observe that by definition of $T(\sigma)$, $T(q) = T(\sigma)$ and therefore

$$\Psi(q) \leq T(\sigma) \leq \frac{1}{2\sigma}.$$

Applying to the extreme terms of this inequality the function φ it follows $q \geq \varphi\left(\frac{1}{2\sigma}\right)$,

$$(2.6) \quad V_q \leq \frac{1}{\varphi\left(\frac{1}{2\sigma}\right)} = \frac{1}{2} U(\sigma).$$

Using (2.1) we obtain

$$(2.7) \quad -\frac{1}{2\sigma} \leq x\alpha - p_1 - \beta \leq \frac{1}{2\sigma} + \frac{1}{2} U(\sigma).$$

10. From (2.2) and (2.7) it follows

$$(2.8) \quad -\frac{1}{2} U(\sigma) \leq x\alpha - \beta - p_1 \leq \frac{3}{2} U(\sigma)$$

and further

$$(2.9) \quad |x\alpha - \beta - p_1| \leq \frac{3}{2} U(\sigma).$$

We see that in order to obtain an approximation of ξ with a error $\leq \varepsilon$, we must make

$$\frac{3}{2} U(\sigma) \leq \varepsilon, \quad \frac{3}{\Psi\left(\frac{1}{2\sigma}\right)} \leq \varepsilon, \quad \Psi\left(\frac{1}{2\sigma}\right) \geq \frac{3}{\varepsilon},$$

$$\frac{1}{2\sigma} \leq \Psi\left(\frac{3}{\varepsilon}\right), \quad \sigma \geq \frac{1}{2 \Psi\left(\frac{3}{\varepsilon}\right)}.$$

Thence $\tilde{\Gamma}(\varepsilon)$ as defined by (1.7) indeed gives a bound for x solving the problem a).

§ 3. A lemma .

11. We prove now the

Lemma. Assume α a real irrational, ω_1, ω_2 positive and a positive integer N . Assume that for any ξ from $0 \leq \xi \leq 1$ there exists a positive integer $x \leq N$ and a convenient integer y such that

$$(3.1) \quad \xi - \omega_1 \leq x\alpha - y \leq \xi + \omega_2.$$

Order all residues $R(x\alpha)$ ($1 \leq x \leq N$) in a monotonically increasing order between 0 and 1 and denote them by

$$r_1 < r_2 < \dots < r_N.$$

Then the lengths of all intervals between two consecutive

r_j as well as the length of the intervals from r_N to $1 + r_1 =: r_{N+1}$

are $\leq \omega_1 + \omega_2$.

12. Proof. Assume that there exists two consecutive r, r_v, r_{v+1} ,

$1 \leq v \leq N$, such that

$$r_{v+1} - r_v > \omega_1 + \omega_2, \quad r_{v+1} - \omega_2 > r_v + \omega_1,$$

and take a ξ such that

$$r_{v+1} - \omega_2 > \xi > r_v + \omega_1.$$

Applying (2.11) to this ξ we obtain a contradiction.

13. Assume on the other hand that the sum of the lengths of the two extreme intervals $\langle 0, r_1 \rangle$ and $\langle r_N, 1 \rangle$ is $> \omega_1 + \omega_2$,

$$r_1 + 1 - r_N > \omega_1 + \omega_2, \quad r_N + \omega_1 < 1 + r_1 - \omega_2.$$

Then there exists a ξ such that

$$r_N + \omega_1 < \xi < 1 + r_1 - \omega_2, \quad r_N < \xi - \omega_1 < \xi + \omega_2 < 1 + r_1.$$

Apply now (3.1) to this ξ ; we obtain

$$r_N < \xi - \omega_1 \leq x\alpha - y \leq \xi + \omega_2 < 1 + r_1,$$

where x, y are integers and $1 \leq x \leq N$. Thence $x\alpha - y$ lies in the open interval $(0, 2)$.

If now $x\alpha - y < 1$ then it must be one of the residues r_x and $r_x > r_N$ is impossible. And if $x\alpha - y$ lies in the open interval $(1, 2)$ then it must be $= 1 + r_x$ and it follows

$$1 + r_x < 1 + r_1, \quad r_x < r_1,$$

which is again impossible. Our lemma is proved;

§ 4. The sums $S(n, \xi)$.

14. Assume α fixed real irrational and ξ real. Put as in (1.17):

$$(4.1) \quad S(n, \xi) := \sum_{v=1}^n R(\xi + v\alpha) \quad . \quad 3)$$

Obviously

$$(4.2) \quad S(n, \xi) = S(n, \xi') \quad (\xi \equiv \xi' \pmod{1}) .$$

Assume ϑ from the half open intervall $\langle 0, 1 \rangle$ and put for an integer v :

$$(4.3) \quad D_v := R(\xi + v\alpha + \vartheta) - R(\xi + v\alpha) .$$

Then

$$(4.4) \quad S(n, \xi + \vartheta) - S(n, \xi) = \sum_{v=1}^n D_v .$$

$$(4.5) \quad D_v \equiv (\xi + v\alpha + \vartheta) - (\xi + v\alpha) \equiv \vartheta \pmod{1} .$$

Consider the intervall

$$(4.6) \quad J := \langle 1 - \xi - \vartheta, 1 - \xi \rangle ,$$

where for $\xi + \vartheta \geq 1$ J is to be understood mod 1, that is to say, consists of the two half open intervalls

$$\langle 2 - \xi - \vartheta, 1 \rangle \quad \text{and} \quad \langle 0, 1 - \xi \rangle .$$

3) This expression was already considered by Sierpinski [1], [2], who, however, only investigated asymptotic properties of $S(n, \xi)$ and did not use the expression (4.9). See also Ostrowski [2] .

I say now that

$$(4.7) \quad D_v = \begin{cases} \varphi - 1 & (R(v\alpha) \in J) \\ \varphi & (R(v\alpha) \notin J) \end{cases}$$

Indeed, if $R(v\alpha) \in J$ then $R(\xi + v\alpha) \in J' := \langle 1 - \varphi, 1 \rangle$ and $R(\xi + v\alpha + \varphi) \in \langle 0, \varphi \rangle$, $R(\xi + v\alpha + \varphi) < \varphi$. But then D_v is a fortiori $< \varphi$ and must have the value $\varphi - 1$.

On the other hand, if $R(v\alpha) \notin J$ then $R(\xi + v\alpha) \notin J' := \langle 1 - \varphi, 1 \rangle$ and $R(\xi + v\alpha + \varphi) \notin \langle 0, \varphi \rangle$. But then $R(\xi + v\alpha + \varphi) > \varphi$ and $D_v > \varphi - 1$, $D_v = \varphi$.

15. Applying the in sec.1 and sec.4 defined expressions $N(n, J)$, $A(n, J)$ to the intervall (4.6) it follows from (4.5), (4.6) and (4.7) :

$$(4.8) \quad S(n, \xi + \varphi) - S(n, \xi) = n\varphi - N(n, J) = A(n, J) .$$

16. Observe that for any couple of natural integers m, n :

$$(4.9) \quad S(n + m, \xi) - S(n, \xi) - S(m, \xi) = S(n, \xi + R(m\alpha)) - S(n, \xi) .$$

Indeed,

$$\begin{aligned} S(n + m, \xi) - S(m, \xi) &= \sum_{v=1}^{n+m} R(\xi + v\alpha) - \sum_{v=1}^m R(\xi + v\alpha) = \\ &= \sum_{v=m+1}^{n+m} R(\xi + v\alpha) = \sum_{v=1}^n R((\xi + m\alpha) + v\alpha) = \\ &= S(n, \xi + m\alpha) = S(n, \xi + R(m\alpha)) . \end{aligned}$$

Define generally for any natural μ :

$$(4.10) \quad J_\mu := \langle 1 - \xi - R(\mu\alpha), 1 - \xi \rangle (\mu \in \mathbb{N}) .$$

Then if we replace in (4.8) ξ with $R(m\alpha)$ the right side expression in (4.9) becomes $A(n, J_m)$ and since the left side expression in (4.9) is symmetric in n and m we obtain

$$(4.11) \quad A(n, J_m) = A(m, J_n) .$$

By definition (1.11) we have

$$|A(n, J)| \leq \hat{A}(n) .$$

Using (4.11) it follows

$$(4.12) \quad |A(n, J_\nu)| = |A(\nu, J_n)| \leq A^*(\nu) .$$

§ 5. Deduction of (1.12) .

17. Let $\sigma \geq 2, [\sigma] =: \sigma_0$ and consider the residues $R(\nu\alpha)$ ($\nu = 1, \dots, \sigma_0$) monotonically ordered:

$$(5.1) \quad 0 < r_1 < \dots < r_{\sigma_0} < 1 .$$

Let $0 < \xi < 1$ and assume first that

$$(5.2) \quad r_1 \leq \xi < r_{\sigma_0} .$$

Then, using (2.8) and the lemma of § 3, we see that

$$r_{v+1} - r_v \leq 2U(\sigma) \quad (v = 1, \dots, \sigma_0 - 1)$$

and it follows that φ lies in one of the intervals $\langle r_v, r_{v+1} \rangle$ ($v = 1, \dots, \sigma_0 - 1$), that is that there exist two $R(v\alpha), \lambda_1, \lambda_2$, such that

$$r_v = \lambda_1 = R(x_1\alpha) \leq \varphi \leq \lambda_2 = r_{v+1} = R(x_2\alpha), \quad (5.3)$$

$$1 \leq x_1 \wedge x_2 \leq \sigma, \lambda_2 - \lambda_1 \leq 2U(\sigma).$$

18. Take an arbitrary but fixed ξ , $0 \leq \xi \leq 1$, and consider the intervals, partly in notation (4.10):

$$\begin{aligned} J_{\lambda_1} &:= \langle 1 - \xi - \lambda_1, 1 - \xi \rangle, \quad J_{\lambda_2} := \langle 1 - \xi - \lambda_2, 1 - \xi \rangle, \\ (5.4) \quad J &:= \langle 1 - \xi - \varphi, 1 - \xi \rangle, \end{aligned}$$

all three with a common end point $1 - \xi$ and of the respective lengths $\lambda_1, \lambda_2, \varphi$. Obviously

$$(5.5) \quad J_{\lambda_1} \subset J \subset J_{\lambda_2}$$

and therefore for any $n \in \mathbb{N}$:

$$N(n, J_{\lambda_1}) \leq N(n, J) \leq N(n, J_{\lambda_2}), \quad n\varphi - N(n, J_{\lambda_2}) \leq n\varphi - N(n, J) \leq n\varphi - N(n, J_{\lambda_1}).$$

Using (1.10) this can be written as

$$(5.6) \quad A(n, J_{\lambda_2}) - n(\lambda_2 - \vartheta) \leq A(n, J) \leq A(n, J_{\lambda_1}) + n(\vartheta - \lambda_1) .$$

Applying (4.12) and using (5.3) it follows from this inequality that

$$(5.7) \quad -(A^*(\sigma) + 2nU(\sigma)) \leq A(n, J) \leq A^*(\sigma) + 2nU(\sigma) ,$$

$$(5.8) \quad |A(n, J)| \leq A^*(\sigma) + 2nU(\sigma) .$$

19. The relation (5.8) has been deduced under the assumption (5.2) and we have now to consider the two remaining intervals for ϑ , $(r_{\sigma_0}, 1)$ and $(0, r_1)$. Assume that

$$(5.9) \quad r_{\sigma_0} < \vartheta < 1 ;$$

then we can still put $\lambda_1 = r_{\sigma_0}$ and it follows $J_{\lambda_1} \subset J$. Using this as in the case (5.2), we obtain

$$(5.10) \quad A(n, J) \leq A^*(\sigma) + 2nU(\sigma) .$$

As to the lower bound for $A(n, J)$, we obtain

$$A(n, J) \geq n\vartheta - H(J) \geq n(\vartheta - 1) ,$$

since obviously $H(n, J) \leq n$. It follows, as by lemma of § 3

$$1 - \vartheta \leq 1 - r_{\sigma_0} \leq 2U(\sigma) :$$

$$-2nU(\sigma) \leq A(n, J) \leq A^*(\sigma) + 2nU(\sigma)$$

and (5.8) follows immediately .

20. Finally, in the case

$$(5.11) \quad 0 \leq \vartheta < r_1$$

we can still take $\lambda_2 = r_1$, $J \in J_{\lambda_2}$ and proceeding again as

in sec. 18 it follows

$$(5.12) \quad A(n, J) \geq -(A^*(\sigma) + 2nU(\sigma)) .$$

As to the upper bound of $A(n, J)$ we have obviously

$$A(n, J) = n\vartheta - N(n, J) \leq n\vartheta \leq nr_1 \leq 2nU(\sigma) \leq A^*(\sigma) + 2nU(\sigma) .$$

(5.8) is now proved for all ϑ , $0 \leq \vartheta < 1$.

Observe now that the interval J as defined in (5.4) can become any partial interval mod 1 of $(0, 1)$, choosing ξ and ϑ conveniently.

It follows therefore from (5.8) :

$$A^*(n) \leq A^*(\sigma) + 2nU(\sigma) .$$

Observe finally that $A^*(x)$ is constant for all x with $R(x) = n$.

We see that we can replace the argument n in A^* by any positive x and the inequality (1.12) is proved.

15

§ 6. Special cases .

21. We assume first that

$$(6.1) \quad T(\sigma) \geq \frac{\gamma}{\sigma}, \quad \gamma > 0.$$

This is the case if the continued fraction development of α has bounded partial quotients, for instance for all quadratic irrationalities. In this case we can take

$$(6.2) \quad \Psi(\sigma) = \frac{\gamma}{\sigma}, \quad \varphi(\sigma) = \frac{\gamma}{\sigma}, \quad U(\sigma) = \frac{1}{\gamma\sigma}.$$

The functional equation (1.12) becomes now,

$$(6.3) \quad A^*(x) \leq A^*(\sigma) + \frac{2}{\gamma} \frac{x}{\sigma}.$$

22. Taking here $\sigma = \frac{x}{e}$ we obtain $A^*(x) \leq A^*\left(\frac{x}{e}\right) + 2e/\gamma$ and generally

$$(6.4) \quad A^*\left(\frac{x}{e^v}\right) \leq A^*\left(\frac{x}{e^{v+1}}\right) + 2e/\gamma \quad (v = 0, 1, \dots).$$

Putting

$$(6.5) \quad n := [\lg x] \leq \lg x < n + 1$$

add the inequality (6.4) over $v = 0, 1, \dots, n$. We obtain

$$A^*(x) \leq A^*\left(\frac{x}{e^{n+1}}\right) + 2(n+1)e/\gamma.$$

If we now put

$$(6.6) \quad \text{Max } A^*(x) = \beta < \infty \quad (0 < x \leq 1)$$

it follows using (6.5)

$$(6.7) \quad A^*(x) \leq 2e(n+1)/\gamma + \beta < \frac{2e}{\gamma} \lg x + 2e/\gamma + \beta .$$

In the particular case of the function $A^*(x)$ defined by (1.10) and (1.11), obviously $\beta = 0$ and we obtain

$$(6.8) \quad A^*(x) \leq 2e(\lg x + 1)/\gamma .$$

23. We consider secondly the case where, assuming two constants g and g' with

$$(6.9) \quad g > g' > 1 ,$$

the function $U(x)$ satisfies for a constant $x_0 \geq 2$ the inequality

$$(6.10) \quad U(x) \leq g' U(gx) \quad (x \geq x_0 \geq 2) .$$

Then we are going to show that

$$(6.11) \quad A^*(x) = O(xU(x)) \quad (x \rightarrow \infty) .$$

More precisely, assuming for an $L > 0$ that

$$(6.12) \quad U(x) \leq L \quad (x_0 \leq x \leq gx_0)$$

and defining D by

$$(6.13) \quad D := \max \left(\frac{A^*(x_0, g)}{x_0 L}, \frac{2g'}{g-g'} \right)$$

we will show that

$$(6.14) \quad A^*(x) \leq DxU(x) \quad (x \geq x_0 \geq 2) \quad .$$

24. In order to prove (6.14) observe first that from the definition (6.13) it follows $D \geq \frac{2g'}{g-g'}$, and therefore

$$(6.15) \quad (D+2)g' \leq gD \quad .$$

Assume now that we have already proved (6.14) for an $x \geq x_0$. Then replacing in (1.12) x by gx and G by x , we obtain from (6.14)

$$A^*(gx) \leq A^*(x) + 2xU(x) \leq (D+2)xU(x) \quad .$$

But this is in virtue of (6.10) and (6.15)

$$\leq (D+2)g'xU(gx) \leq DgxU(gx) \quad .$$

We see that (6.14) is true for any gx whenever it is true for x . We have therefore only to prove (6.14) throughout the interval $(x_0 \leq x \leq gx_0)$. This inequality follows, since $A^*(x)$ is monotonically increasing, by (6.12) and (6.13) from

$$A^*(gx_0) \leq Dx_0 L \quad .$$

The relation (6.14) is proved.

25. The simplest special case is that of

$$(6.16) \quad \Psi(x) = \frac{c}{x^{\frac{1}{\xi}}} \quad (\xi > 1) \quad .$$

Then we have

$$\varphi(y) = \left(\frac{c}{y}\right)^{\frac{1}{\xi}} , \quad \varphi\left(\frac{1}{2y}\right) = (2cy)^{\frac{1}{\xi}} , \quad U(x) = \frac{2}{(2cx)^{\frac{1}{\xi}}} ,$$

$$(6.17) \quad \Lambda^*(x) = O(x^{1-\frac{1}{\xi}}) \quad (x \rightarrow \infty) \quad .$$

In particular it follows, in notations of § 1,

$$(6.18) \quad N(x, J) = x(|J| + O(\frac{1}{x^{\frac{1}{\xi}}})) \quad (x \rightarrow \infty) .$$

As to the detailed investigation of the case considered in sec. 23 - 24 we will give it in another paper dealing with multi-dimensional approximations.

Bibliography

- [1] Behnke, H. , Zur Theorie der Diophantischen Approximationen, Hamb. Math. Abh. , 3, 1924, pp. 261-318.
- [1] Bohl, P. , Über ein in der Theorie der säkularen Störungen vorkommendes Problem, Crelles Journal, vol.135, 1909, pp.222-224. Reprinted (in russian translation) in P.Bohl's collected papers, Riga, 1974, pp.335-336
- [1] Cassels, J.W.S. , An Introduction to Diophantine Approximation, Cambridge Tracts, Cambridge University Press, 1951.
- [1] Hardy, G.H. and Littlewood, J. E. , Some problems of Diophantine approximation. The lattice points of a right-angled triangle.Proc. L.M.S.20(2), pp.15-36,(1921).Reprinted in Hardy's collected papers, vol.1,pp.136-157.
- [2] Hardy, G.H. and Littlewood, J.E. , Some problems of Diophantine approximation. The lattice points of a right-angled triangle (second memoir). Hamb.Math.Abh.,vol.1,1922,pp.212-249. Reprinted in Hardy's collected papers, vol.1,pp.159-196.
- [1] Hecke, E. , Über analytische Funktionen und die Verteilung von Zahlen mod eins, Hamb.Math.Abh.,1 (1921), pp.54-76, in particular pp.72-76.Reprinted in Hecke's Math. Werke, Göttingen 1959,pp.313-335, in particular § 6, pp.331-335.
- [1] Khintchine,A.J. , The qualitative conception of Kronecker's approximation theory (in russian), Isv.Akad.Sz.USSR,math.,12 (1948) pp. 113-122.
- Littlewood, J.E. , see Hardy, G.H. and Littlewood, J.E.
- [1] Ostrowski, A. , Bemerkungen zur Theorie der Diophantischen Approximationen (I - III). Aus Mitteilungen an Herrn Hecke. Hamb.Math.Abh.,1,1921,pp.77-98,250.
- [2] Ostrowski, A. , Notiz zur Theorie der Diophantischen Approximationen Jahresbericht d.D.M.V.,XXXVI,1927, pp.178-180.
- [3] Ostrowski, A. , Zur Theorie der linearen Diophantischen Approximationen. Jahresbericht d.D.M.V.,XXXIX,1930,pp.34-46.
- [1] Sierpinski, W. , Sur les nombres irrationnels, Krakau.Anz.,1909(2), pp.725-727.
- [2] Sierpinski, W. , Sur la valeur asymptotique d'une certaine somme. Krakau.Anz.(A),1910(2), pp.9-11.
- [1] Tchebyshev, P.L. , On an arithmetical problem (in russian), Notices S.Petersburg Acad.Sciences, vol.X, 1866, reprinted in french translation in Oeuvres, vol.I,1899, pp.639-684.
- [1] Weyl, H. , Über die Gibbsche Erscheinung und verwandte Konvergenzphänomene.Rend.Circ.Mat.Palermo, 30, 1910, pp.377-407, in particular p.406. Reprinted in Weyl's collected papers, vol.1, Springer 1968, pp.352-353.

September 1977

BIM 51

BASEL MATHEMATICAL NOTES

ON THE ERROR TERM IN MULTIDIMENSIONAL
DIOPHANTINE APPROXIMATION

by A.M.Ostrowski

Sponsored in part by the Swiss
National Science Foundation.

Sponsored in part under the Grant
DA-ERO-75-6-035 of the European
Research Office, United States Army,
to the Institute of Mathematics,
University of Basel.

§ 1. Introduction

1. Consider the euclidian space R^m and m real irrational numbers α_μ ($\mu=1, \dots, m$) such that $\alpha_1, \dots, \alpha_m, 1$ are linearly independent with respect to \mathbb{Z} , that is that for integers $\varepsilon_1, \dots, \varepsilon_m, \varepsilon_0$ the relation

$$(1.1) \quad \varepsilon_1 \alpha_1 + \dots + \varepsilon_m \alpha_m + \varepsilon_0 = 0$$

only holds if all ε_μ vanish.

Denote generally for any real a by the symbol $\|a\|$ the distance of a from the nearest integer. Further, consider a variable m vector, z , with integer components, $z = (z_1, \dots, z_m)$, which is assumed never to vanish and put

$$\|z\|_\infty := \max_{\mu} |z_\mu|.$$

Then we can use as the "measure of independence" of the α_μ any continuous and strictly monotonically decreasing $\Psi(\sigma) \downarrow 0$ ($\sigma \gg 1$) such that

$$(1.2) \quad \Psi(\sigma) \leq \min \left\| \sum_{\mu=1}^m z_\mu \alpha_\mu \right\| \quad (1 \leq \|z\|_\infty \leq \sigma), \quad (\sigma \gg 1).$$

2. By $P(k \alpha_\mu)$ we denote generally the point

$$P(k \alpha_\mu) : (k \alpha_1, \dots, k \alpha_m) \bmod 1.$$

We define a "proper interval" J in R^m as a cartesian product of m linear segments, open in the direction of increasing coordinates,

$$a_\mu \leq x_\mu < b_\mu \quad (\mu = 1, \dots, m).$$

Then the volume of J , $|J|$, is the product of the lengths of these segments. If an interval is considered mod 1, we take generally two points of \mathbb{R}^m as identical if their corresponding coordinates differ by integers. Then usually the points of J have to be taken with a convenient multiplicity. If all points of J , taken mod 1, are simple, J can be considered as a part of the unity cube, $0 \leq x_\mu < 1$ ($\mu = 1, \dots, m$) and is called simple. Then $|J|$ is defined as cartesian product of m segments mod 1, $0 \leq a_\mu \leq x_\mu < b_\mu \leq 1$ or $(0 \leq x < b_\mu) \cup (a_\mu \leq x < 1)$. Denoting by δ_μ in the first case 0 and in the second case 1, the length of the μ -th edge of J is

$$(1.3) \quad b_\mu - a_\mu + \delta_\mu,$$

while the volume of J becomes

$$|J| = \prod_{\mu=1}^m (b_\mu - a_\mu + \delta_\mu).$$

We put generally

$$\tau(J) := \min_{1 \leq \mu \leq m} (b_\mu - a_\mu + \delta_\mu).$$

3. The essential point of Kronecker's theory of irrationals is the result that for any J and for at least one integer q :

$$(1.4) \quad P(q\alpha) \in J.$$

This result was sharpened by Weyl [1] who proved that for a given $\varepsilon > 0$ for any J with $\tau(J) \leq \varepsilon$ the integer q in (1.4) can be chosen $\leq y(\varepsilon)$, with a $q(\varepsilon)$ independent of the special J with $\tau(J) \leq \varepsilon$:

$$(1.5) \quad P(q_J \alpha) \in J, \quad q_J \leq y(\varepsilon).$$

4. We will denote by $N(x, J)$ the number of all $P(v\alpha)$ with $1 \leq v \leq x$ lying mod 1 in J ,

$$(1.6) \quad N(x, J) := N(\nu : P(\nu\alpha) \in J, 1 \leq \nu \leq x) .$$

Then it follows immediately from Weyl's theorem (1.5), that for any fixed simple intervall $J \bmod 1$ the relation holds

$$(1.7) \quad N(x, J) = x |J| + o(x) \quad (x \rightarrow \infty) .$$

§. We introduce, for $x \geq 1$, $A(x)$ by

$$(1.8) \quad A(x) := \sup_{1 \leq \nu \leq x} |N(x, J) - x|J|| ,$$

where J runs through all simple intervalls mod 1 in the unity cube in R^m , and denote by $\sigma = \varphi(\tau)$ the inverse function of $\varphi = \Psi(\sigma)$ in (1.2). We put further

$$(1.9) \quad \gamma_0 := 2^{m+1}/((m+1)!)^2, \quad \xi(x) := \frac{1}{\gamma_0 \varphi(\frac{1}{\gamma_0 x})}$$

6. It is easily seen that $A(x) = o(x) \quad (x \rightarrow \infty)$. Our aim is to improve this estimate using the function $\xi(y)$. Our essential result is the Theorem 1. $A(x)$, as defined by (1.8), satisfies for $m > 1$ an inequality

$$(1.10) \quad A(x) \leq \alpha A(y) + \beta x \xi(y), \quad 1 \leq y \leq x, \quad \alpha > 1, \beta > 0$$

with constants $\alpha \geq 1$ and $\beta > 0$ depending only on m .

This theorem is proved in the §§ 2 - 5 while the inequality (1.10) is discussed in § 6 under different assumptions about $\xi(y)$. In the case $m = 1$ an inequality (1.10) holds even with $\alpha = 1$. This case however has been already discussed in Ostrowski [2] .

7. Before attacking the problem of $o(x)$ in (1.7) we have to obtain a relationship between ξ and $y(\xi)$ in (1.5) and $\Psi(\sigma)$ in (1.2). This relationship follows in a particularly simple and fundamental way from a special case of an important theorem due to Khintchine (Khintchine [1]¹⁾. We obtain from this theorem for the constant γ_0 from (1.9) the relation (sec.13, (2.10)):

1) My attention was drawn kindly to this theorem by J.W.S. Cassels (Cassels [1] , pp. 97-99) .

(1.11)

$$y = \frac{1}{\psi_0 \psi\left(\frac{1}{y_0 \varepsilon}\right)} .$$

8. It is well known since Dirichlet that $\psi(y) = O\left(\frac{1}{y^m}\right)$. If there exists a ϑ , $0 < \vartheta < 1$, such that

$$(1.12) \quad \psi(y) \geq \frac{c}{y^{\vartheta}} \quad (y \geq y_0)$$

then we show that (sec. 33)

$$(1.13) \quad \Lambda(y) = O(x^{1-\vartheta}) .$$

This is in particular always the case if the $\alpha_1, \dots, \alpha_m$ are algebraic. However the estimate of $\Lambda(x)$ with the exponent $1 - \vartheta$ could only be obtained using Khintchine's theorem, published 1949. In Ostrowski [1], 1930, we used a weaker result than (1.13), due to Landau.

In the case $m = 1$ the fact that (1.13) follows from (1.12) has been already proved by Hecke 1922, however with a method which apparently cannot be generalized to $m > 1$. 1930 we announced the results corresponding to (1.12) and (1.13), however in the form $\Lambda(x) = O(x^{\alpha})$, $0 < \alpha < 1$, (see Ostrowski [2]).

9. More generally assume $k(y)$ as a positive constant or a continuous positive function strictly increasing to ∞ , such that

$$\psi(y) \geq \frac{c}{y^{\vartheta} k((Dy)^{\vartheta})} \quad (y \geq y_0), \quad c \wedge D > 0, \text{ const.}$$

Then for a conveniently defined $\ell(x)$:

$$(1.14) \quad \Lambda(x) = O(x^{1-\vartheta} \ell(x)^{\vartheta}) .$$

This follows from lemma 5, sec 43.

$$\text{If in particular } x k'(x) = o\left(\frac{k(x)}{\lg k(x)}\right)$$

then we obtain even

$$(1.15) \quad \Lambda(x) = O(x^{1-\vartheta} k(x)^{\vartheta}) \quad (\text{Theorem 2, sec 50}).$$

10. Our proof of (1.10) was given for $m=1$, 1930, in Ostrowski [1]. Its essential point was our lemma 3 of § 3 which we developed l.c. for $m=1$, but indicated that the whole discussion can be generalized to $m>1$. In the mean time, 1950, S. Hartmann (Hartmann [1]) has developed in a very careful way the corresponding generalization of the lemma 3 to $m>1$, discussing also the limiting cases. As we need only a part of this argument, we give in § 3 our original proof, which is a straightforward generalisation of that given for $m=1$ in Ostrowski [1].

§ 2. Use of Khintchine's lemma

11. We formulate first one part of Khintchine's theorem in the form in which it was given by Cassels ([1] p.99), but changing conveniently the notation. We will denote generally for an n -vector, $\xi = (x_1, \dots, x_n)$, by $|\xi|_\infty$ the norm $|\xi|_\infty := \max |x_v|$.

Lemma 1. Let m and n be natural integers with $m+n = \ell$. Consider a real $(n \times m)$ -matrix, $A = (\alpha_{\nu\mu})$, $\mu = 1, \dots, m$; $\nu = 1, \dots, n$ and the linear forms

$$(2.1) \quad H_\mu(\xi) = \sum_{\nu=1}^n \alpha_{\nu\mu} x_\nu \quad (\mu = 1, \dots, m),$$

$$(2.2) \quad H_\nu(\xi) = \sum_{\mu=1}^m \alpha_{\nu\mu} z_\mu \quad (\nu = 1, \dots, n),$$

where the x_ν and the z_μ are respectively the components of the n -vector ξ and the m -vectors ζ . Consider two positive constants ε, γ and a real m -vector β with components b_1, \dots, b_m .

Then, in order that there exists an integral vector ξ satisfying the relations

$$(2.3) \quad |H_\mu(\xi) - b_\mu| \leq \varepsilon \quad (\mu = 1, \dots, m), \quad |\xi|_\infty \leq \gamma$$

it is sufficient that for $\gamma := 2^{\ell-1}/(\ell!)^2$ the following relation holds for every integer m -vector ζ :

$$(2.4) \quad \left\| \sum_{\mu=1}^m b_{\mu} z_{\mu} \right\| \leq \gamma \max(y \max \|N_{\nu}(\zeta)\|, \varepsilon |\zeta|_{\infty}) .$$

As a matter of fact the complete formulation of Khintchine's theorem contains also the necessary condition for (2.3), which we however do not need .

12. For our purpose we must now specialise the assumptions of Khintchine's theorem.

Assume $n=1$, $\ell=m+1$ and observe that $\gamma_0 := 2^{m+1} / [(m+1)!]^2 = \frac{4}{\pi^2}$. The n -vector ξ becomes a scalar which we will denote by q , the elements of the matrix A become $\alpha_{\mu\nu} := \alpha_{\mu}$, so that $H_{\mu}(\xi)$ becomes $q\alpha_{\mu}$ and the linear forms $N_{\nu}(\zeta)$ become $H(\zeta) = \sum_{\mu=1}^m \alpha_{\mu} z_{\mu}$. The requirements (2.3) of Khintchine's theorem become

$$(2.5) \quad \|q\alpha_{\mu} - b_{\mu}\| \leq \varepsilon \quad (\mu=1, \dots, m), \quad |q| \leq \gamma .$$

It follows then from the condition (2.4) of Khintchine's theorem that (2.5) can be certainly realized by a rational integer q if for any m -vector ζ we have

$$(2.6) \quad \left\| \sum_{\mu} b_{\mu} z_{\mu} \right\| \leq \frac{\gamma}{2} \max[\gamma \|H(\zeta)\|, \varepsilon |\zeta|_{\infty}], \quad \gamma_0 = 2^{m+1} / ((m+1)!)^2 .$$

13. The condition (2.6) is sharpened replacing the left side expression by $1/2$. As it is certainly satisfied if $\gamma_0 \varepsilon |\xi|_{\infty} > 1$, it suffices to consider ξ with

$$(2.7) \quad \varepsilon |\zeta|_{\infty} \leq \frac{1}{\gamma_0} .$$

Thence our condition becomes:

$$(2.8) \quad 1 \leq \gamma_0 \gamma \|H(\zeta)\| \text{ follows always from } |\zeta|_{\infty} \leq \frac{1}{\gamma_0 \varepsilon} .$$

If we now assume that (1.1) holds and use the definition (1.2) of $\Psi(t)$, (2.3) is satisfied if

$$\Psi\left(\frac{1}{y_0 \varepsilon}\right) \geq \frac{1}{y_0 y},$$

and we can take y in (2.5) as

$$(2.9) \quad y = y(\varepsilon) := \frac{1}{y_0 \Psi\left(\frac{1}{y_0 \varepsilon}\right)}.$$

Using the inverse function to Ψ, φ , it follows as in (1.9)

$$(2.10) \quad \frac{1}{y_0 \varepsilon} = \varphi\left(\frac{1}{y_0 \varepsilon}\right), \quad \varepsilon = \varepsilon(y) := \frac{1}{y_0 \varphi\left(\frac{1}{y_0 y}\right)}.$$

Lemma 2. For any $y \geq \frac{1}{y_0 \Psi(1)}$ there exists an integer q with

$$(2.11) \quad |q| \leq y, \quad \|a_\mu - b_\mu\| \leq \varepsilon(y) = \frac{1}{y_0 \varphi\left(\frac{1}{y_0 y}\right)} \quad (\mu = 1, \dots, n).$$

§ 3. A lemma

14. In what follows we will consider a sequence

$$(3.1) \quad P(v \alpha_\mu) \quad (v = 1, \dots, n)$$

for a fixed integer $n \geq 1$. We define the symbol $[a - 0]$ as $[a]$ if a is not integer and $a - 1$ if a is integer.

15. Lemma 3. Consider a simple interval $J_0 \bmod 1$ contained in the unity cube, as characterized in sec. 2, and assume the α_μ as in sec. 1. Then there exist two intervals J'_0 and $J''_0 \bmod 1$ in \mathbb{R}^n , obtained from J_0 by parallel translations, such that

$$(3.2) \quad N(n, J'_0) \leq n |J_0|, \quad N(n, J''_0) \geq n |J_0|.$$

16. Proof. Without loss of generality we can assume, that all a_μ lie in the open interval $(0,1)$ and further, that J is not identical with the unit cube, but "begins" at the origine, that is that all a_μ in (3.1) vanish. Denote the length of the μ -edge of J_0 by d_μ , where

$$0 < d_\mu \leq 1 \quad (\mu=1, \dots, m), \quad |J_0| = d_1 \cdots d_m < 1.$$

17. If we shift J_0 in the directions of the x_μ by the integers q_μ , we obtain a proper interval which will be denoted by J_{q_1, \dots, q_m} .

Then the original J_0 can be written as $J_{0, \dots, 0}$. Obviously J_{q_1, \dots, q_m} is the cartesian product of the segments q_1, \dots, q_m

$$[q_\mu d_\mu, (q_\mu + 1)d_\mu) \quad (\mu=1, \dots, m) \quad .$$

18. We let now, for positive integers Q_1, \dots, Q_m , run each q_μ through $0, 1, \dots, Q_\mu - 1$. Then all intervals obtained in this way form together an interval J^* with the edges $Q_\mu d_\mu$ ($\mu=1, \dots, m$) and its volume is

$$Q |J_0| = \prod_{\mu=1}^m (Q_\mu d_\mu), \quad Q := Q_1 \cdots Q_m \quad .$$

Put further

$$(3.3) \quad N_{q_1, \dots, q_m} := N(n, J_{q_1, \dots, q_m})$$

and denote by N the sum of all N_{q_1, \dots, q_m} ($q=0, 1, \dots, Q_1; \dots; q_m=0, 1, \dots, Q_m$):

$$(3.4) \quad N := \sum_{q_1, \dots, q_m} N_{q_1, \dots, q_m}$$

19. Denote by $f(v)$ the number of all points in J^* which, considered mod 1, coincide with the $P(v\alpha_\mu)$ from (9.1). These points have the coordinates

$$[\varepsilon_1 + R(v\alpha_1), \varepsilon_2 + R(v\alpha_2), \dots, \varepsilon_m + R(v\alpha_m)]$$

where

$$0 \leq \varepsilon_1 < [\gamma_1 d_1 - R(v\alpha_1) - 0], \dots, 0 \leq \varepsilon_m < [\gamma_m d_m - R(v\alpha_m) - 0].$$

Therefore we have

$$(3.5) \quad f(v) = \prod_{\mu=1}^m [\gamma_\mu d_\mu + 1 - R(v\alpha_\mu) - 0].$$

By summation over $v = 1, \dots, n$ we obtain the number of all points in J^* equal mod 1 to the points (3.1), that is N . Dividing by Q we obtain finally

$$(3.6) \quad \frac{1}{Q} \sum_{v=1}^n \prod_{\mu=1}^m [\gamma_\mu d_\mu + 1 - R(v\alpha_\mu) - 0] = \sum_{q_1, \dots, q_m}^N q_1, \dots, q_m.$$

20. If we let all γ_μ increase to ∞ , the left side expression in (3.6) tends to $n |J_0|$. Therefore the same holds in the right hand expression and we obtain

$$(3.7) \quad \sum_{q_1, \dots, q_m}^N q_1, \dots, q_m \rightarrow n |J_0| \quad (\forall \gamma_\mu \rightarrow \infty).$$

21. Assume first that $n |J_0|$ is not integer and lies between $k, k+1$. Then obviously, as soon as the left side expression in (3.7) lies in the open interval $(k, k+1)$, it is impossible that all N_{q_1, \dots, q_m}

in (3.7) are $\geq k+1$. Neither can all these N_{q_1, \dots, q_m} be $\leq k$.

Therefore there exist at least two different J_{q_1, \dots, q_m} , say J'_0, J''_0 ,

so that (3.2) is satisfied.

22. Assume now that $n | J_0 |$ is an integer. If there exist two J_{q_1}, \dots, J_{q_m} , say J'_0 and J''_0 , so that

$$N(n, J'_0) < n | J |, \quad N(n, J''_0) > n | J_0 |,$$

(3.2) is again satisfied. Otherwise for all J_{q_1}, \dots, J_{q_m} the corresponding N_{q_1}, \dots, N_{q_m} in (3.7) are equal to $n | J_0 |$ and then we can take $J'_0 = J''_0 = J_0$, and the relations (3.2) are satisfied with the equality sign.

§ 4. An upper limit for $N(n, J) - n | J |$.

23. Consider a simple interval $J \bmod 1$ with the edges d_1, \dots, d_m and a positive $\varepsilon < 1/4$. Assume first that

$$(4.1) \quad 1 - d_\mu > 2\varepsilon \quad (\mu=1, \dots, m).$$

Let J_0 be an interval concentric to J with the edges $d_1 + 2\varepsilon, \dots, d_m + 2\varepsilon$.

(See fig. 1, p. 28, for $m = 2$).

By (4.1) J_0 is also a simple interval and we have

$$(4.2) \quad |J_0| = \prod_{\mu=1}^m (d_\mu + 2\varepsilon), \quad |J| = \prod_{\mu=1}^m d_\mu.$$

24. By the first inequality (3.2) there exists an interval J'_0 congruent with J_0 such that

$$(4.3) \quad N(n, J'_0) \leq n|J_0|.$$

Consider a cube, C , with the edges parallel to the axes and of the length ε , placed so that it has with J only one point of the boundary in common, the vertex $E = (e_1, \dots, e_m)$, and lies completely in J_0 (see the hatched square C in the fig. 1, p. 28).

25. Consider the vertex of J_0 corresponding to E and the corresponding vertex of J'_0 , $E' = (e'_1, \dots, e'_m)$. Then by what has been proved in sec. 13 about the relation (1.5) it follows that for a convenient positive integer $q \in \mathcal{V}(\varepsilon) = 1/\left[\gamma_0 \psi\left(\frac{4}{\gamma_0 \varepsilon}\right)\right]$ and convenient integers q_1, \dots, q_m the relations hold

$$(4.4) \quad e_\mu - e'_\mu = R(q\alpha_\mu) + q_\mu \varepsilon, \quad 0 < |q_\mu| < 1 \quad (\mu = 1, \dots, m).$$

It follows that if we apply the parallel translation first by the vector $q\alpha$ and then by the integer vector (q_μ) to the interval J'_0 , this interval goes over into a congruent interval \bar{J} , which has the property that the vertex of \bar{J} corresponding to E'_0 lies in the hatched domain C . Obviously, J is contained in \bar{J} .

26. Consider the translation from J'_0 to \bar{J} . To the points $P(v\alpha)$, $1 \leq v \leq n$, lying in J'_0 correspond the points congruent mod 1 to $P((v+q)\alpha)$ that is to the points $P(v\alpha)$, $q+1 \leq v \leq n+q$. Their number is

$$N(n, J_0) = N(n+q, \bar{J}) - N(q, \bar{J}).$$

But the minuend here can be written as

$$N(n+q, \bar{J}) = N(n, \bar{J}) + N(R(v\alpha), R(v\alpha) \in \bar{J}, n+1 \leq v \leq n+q).$$

Here the last summand can be again written as $N(q, J^*)$ if we denote by J^* the interval obtained from \bar{J} by the parallel translation with the vector $R(n\alpha)$. And obviously $|J^*| = |J_0| = |\bar{J}|$.

27. We obtain therefore from (4.3)

$$(4.5) \quad N(n, J'_0) = N(n, \bar{J}) - N(q, \bar{J}) + N(q, J^*) \leq n|J_0|,$$

$$N(n, \bar{J}) \leq n|J_0| + N(q, \bar{J}) - N(q, J^*).$$

As $J \in \bar{J}$, it follows further

$$N(n, J) \leq N(n, \bar{J}) \leq n|J_0| - N(q, \bar{J}) + N(q, J^*) .$$

Since $|J| = |J^*|$, we can write this in the form

$$N(n, J) \leq n|J_0| - \left[N(q, \bar{J}) - n|\bar{J}| \right] + \left[N(q, J^*) - n|\bar{J}| \right] .$$

The two last bracket terms on the right have moduli $\leq \Lambda(q)$ and we obtain further

$$N(n, J) - n|J_0| \leq 2\Lambda(q) .$$

On the other hand it follows from (4.2) :

$$|J_0| - |J| \leq \prod_{\mu=1}^m (d_\mu + 2\varepsilon) - \prod_{\mu=1}^m d_\mu \leq (1 + 2\varepsilon)^m - 1 < 2^{m+1} \varepsilon ,$$

as $2\varepsilon < 1$ and the development of $|J_0| - |J|$ in products of the d_μ has positive coefficients .

Since $q \leq y(\varepsilon)$ and $\Lambda(q)$ is not decreasing we can finally write

$$(4.6) \quad N(n, J) - n|J| \leq 2\Lambda(y(\varepsilon)) + 2^{m+1} n\varepsilon .$$

28. (4.6) has been derived assuming the condition (4.1) .

If this condition is not satisfied , we can by halving each edge of the unity cube decompose the unity cube into the sum of 2^m cubes of the edge length $1/2$. Correspondingly J is decomposed into at the most 2^m intervals $J^{(\nu)}$ ($\nu = 1, 2, \dots$) with the edge lengths $\leq \frac{1}{2}$.

For each of the intervals $J^{(\nu)}$ the condition (4.1) is satisfied so that we can write

$$N(n, J^{(\nu)}) - n|J^{(\nu)}| \leq 2\Lambda(y(\varepsilon)) + 2^{m+1} n\varepsilon .$$

Summing over ν it follows

$$(4.7) \quad N(n, J) - n|J| \leq 2^{m+1} \Lambda(y(\varepsilon)) + 2^{2m+1} n\varepsilon .$$

§ 5. A lower limit for $N(n, J) - n|J|$.

29. We consider again the simple interval J mod 1 of the sec.23 with the edges d_1, \dots, d_m , but assume first that for a positive $\varepsilon < 1/4$ the relations hold:

$$(5.1) \quad d_\mu > 2\varepsilon \quad (\mu=1, \dots, m) .$$

Let now J_0 be an interval concentric to J with the edges $d_1 - 2\varepsilon, \dots, d_m - 2\varepsilon$. J_0 is again a simple interval with

$$(5.2) \quad |J_0| = \prod_{\mu=1}^m (d_\mu - 2\varepsilon) , \quad |J| = \prod_{\mu=1}^m d_\mu .$$

By the second inequality in (3.2) there exists an interval J'_0 congruent with J_0 and such that

$$(5.3) \quad N(n, J'_0) \geq n|J_0| .$$

Consider a cube, C , with the edges parallel to the axes and of the length ε which has with J_0 only an edge $E = (e_1, \dots, e_m)$ in common (see the hatched square in fig.2, p.28).

30. Consider the vertex of J'_0 , $E' = (e'_1, \dots, e'_m)$, which corresponds to E . Then, by what has been proved in sec.13 about the relation (1.3), it follows that for a convenient positive integer $q \leq \gamma(\varepsilon)$ and convenient integers q_1, \dots, q_m the relations hold:

$$e_\mu - e'_\mu = R(q\varepsilon) + q_\mu + \theta_\mu \varepsilon , \quad 0 \leq \theta_\mu < 1 \quad (\mu=1, \dots, m).$$

We see that if we apply to the interval J'_0 the parallel translations first by the vector $q\varepsilon$ and then by the integer vector $q = (q_\mu)$, this interval goes over into a congruent interval \bar{J}

which has the property that the vertex of \bar{J} corresponding to E'_0 lies in the hatched domain C. Obviously J contains \bar{J} (see fig.2 for $m=2$).

31. By parallel translation from J'_0 to \bar{J} , to the points $P(v\alpha)$, $1 \leq v \leq n$, lying in J'_0 correspond points congruent mod 1 to $P((v+q)\alpha)$ that is to the $P(v\alpha)$, $q+1 \leq v \leq n+q$.

Their number is

$$N(n, \bar{J}) = N(n+q, \bar{J}) - N(q, \bar{J}).$$

But the minuend can be written as

$$N(n+q, \bar{J}) = N(n, \bar{J}) + N(R(v\alpha) \in \bar{J}, n+1 \leq v \leq n+q).$$

Here the last summand can be again written as $N(q, J^*)$ if we denote by J^* the interval obtained from \bar{J} by the parallel translation with the vector $R(n\alpha)$. And obviously $|J^*| = |J'_0| = |\bar{J}|$.

We obtain therefore

$$N(n, J'_0) = N(n, \bar{J}) - N(q, \bar{J}) + N(q, J^*),$$

$$N(n, \bar{J}) = N(n, J'_0) + [N(q, \bar{J}) - n|\bar{J}|] - [N(q, J^*) - n|\bar{J}|].$$

But here $N(n, \bar{J})$ is $\leq N(n, J)$ while $N(n, J'_0)$ is, by (5.3), $\geq n|J'_0|$. As both bracket expressions are $\leq A(q)$ we obtain

$$(5.4) \quad N(n, J) \geq n|J'_0| - 2A(q).$$

32. On the other hand, similarly as in sec.27,

$$|J'_0| = |J'_0| = \prod_{\mu=1}^m (d_\mu - 2\varepsilon) \geq \prod_{\mu=1}^m d_\mu - 2\varepsilon 2^m = |J| - 2^{m+1}\varepsilon.$$

Introducing this into (5.4) we obtain finally

$$(5.5) \quad N(n, J) - n|J| \geq -(2A(q) + 2^{m+1}n\varepsilon).$$

33. If we now drop the restriction (5.1) and assume that at least one of the d_μ is $\leq 2\varepsilon$, obviously $|J| \leq 2\varepsilon$. But then the relation (5.5) holds again and is therefore now proved independently of the restriction (5.1).

Combining (4.7) and (5.5) it follows

$$|N(n, J) - n|J|| \leq 2^{m+1} \Lambda(y(\varepsilon)) + 2^{2m+1} n \varepsilon.$$

Referring now to the definition (1.8) of $\Lambda(x)$ it follows now since obviously $\Lambda(x) = \Lambda([x])$, while $y(\varepsilon)$ in (2.10) is continuous,

$$(5.6) \quad \Lambda(x) \leq 2^{m+1} \Lambda(y) + 2^{2m+1} x \varepsilon(y) \quad (x \wedge y \geq 1) \dots$$

This functional inequality the derivation of which is the essential point of our method, is a special case of the following inequality

$$(5.6) \quad \Lambda(x) \leq \alpha \Lambda(y) + \beta x \varepsilon(y) \quad (\alpha \geq 1, \beta > 0, x \wedge y \geq 1),$$

where α and β are given constants.

§ 6. Discussion of the fundamental inequality.

34. We are going first to treat the general inequality (5.6). We assume generally about $\varepsilon(y)$ that it is positive and monotonically decreasing to 0 with $y \rightarrow \infty$ while $\Lambda(x)$ is assumed to be positive and monotonically increasing for $x \geq 1$.

Lemma 4. Assume that for four constants ε, g', L and x_0 with

$$(6.1) \quad \varepsilon > 1, \quad 0 < g' < \frac{\varepsilon}{\alpha}, \quad L > 0, \quad x_0 \geq 1$$

the following relations are satisfied

$$(6.2) \quad \varepsilon(x) \leq g' \varepsilon(gx) \quad (x_0 \leq x),$$

$$(6.3) \quad \varepsilon(x) \geq L \quad (1 \leq x \leq x_0).$$

Under these conditions

$$(6.4) \quad A(x) = O(x \varepsilon(x)) \quad (x \uparrow \infty)$$

and more precisely

$$(6.5) \quad A(x) \leq Dx \varepsilon(x) \quad (x \geq 1),$$

where D is defined by

$$(6.6) \quad D := \max \left(\frac{A(x_0)}{L}, \frac{g g' \beta}{g - g' \alpha} \right).$$

35. Proof. If we first assume that $1 \leq x \leq x_0$ it follows by (6.3) and (6.6) as $A(x)$ is increasing,

$$A(x) \leq A(x_0) \leq DL \leq Dx \varepsilon(x)$$

and we see that (6.5) holds for $1 \leq x \leq x_0$.

It is therefore sufficient to prove that, if (6.5) holds for an $x \geq 1$ it also holds for gx . But replacing in (5.6) x with gx and y with x it follows

$$A(gx) \leq \alpha A(x) + \beta gx \varepsilon(x) = (\alpha D + \beta g)x \varepsilon(x) \leq \frac{g}{g'} (\alpha D + \beta g)(gx) \varepsilon(gx)$$

and here the factor $\frac{g}{g'} (\alpha D + \beta g)$ is $\leq D$ as follows immediately from

$$D \geq \frac{g g' \beta}{g - g' \alpha}.$$

The inequality (6.5) is completely proved.

36. Consider now instead of the assumptions made in sec. 34 the assumptions

$$(6.7) \quad \alpha > 1, \quad f(x) = \frac{f(x)}{x}, \quad 0 < f(x) < 1, \quad \Lambda(x) \leq x \quad (x \geq 1).$$

Then choosing

$$(6.8) \quad l_{\alpha} z := \sqrt{l_{\alpha} \alpha \lg x}$$

put

$$(6.9) \quad n := \left\lceil \frac{l_{\alpha} x}{l_{\alpha} z} \right\rceil = \left\lceil \sqrt{\frac{l_{\alpha} x}{l_{\alpha} \alpha}} \right\rceil.$$

It follows $1 \leq x z^{-n} < z$,

$$(6.10) \quad \alpha^n \leq e^{\frac{l_{\alpha} x}{l_{\alpha} z}} = e^{\sqrt{l_{\alpha} \alpha \lg x}} = z.$$

37. Put in (5.6) $y = \frac{x}{z}$. We obtain, by (6.7),

$$\Lambda(x) \leq \alpha \Lambda\left(\frac{x}{z}\right) + \beta z.$$

Writing this inequality for x/z^v instead of x and multiplying it with α^v ,

$$\alpha^v \Lambda\left(\frac{x}{z^v}\right) \leq \alpha^{v+1} \Lambda\left(\frac{x}{z^{v+1}}\right) + \beta z \alpha^v.$$

Summing over $v=0, 1, \dots, n-1$ it follows

$$\Lambda(x) \leq \alpha^n \Lambda\left(\frac{x}{z^n}\right) + \beta z \sum_{v=0}^{n-1} \alpha^v.$$

We use now (6.10) and obtain

$$(6.11) \quad \Lambda(x) \leq x^2 \left(1 + \frac{\beta}{\alpha-1} \right) = o(e^{\sqrt{\lg \alpha \lg x}}).$$

The reader may be reminded that in the case $\alpha = 1$, from $\tau(x) = O(1/x)$ follows

$$\Lambda(x) = O(\lg x)$$

as is shown in Ostrowski [2].

38. Replace now the conditions of sec. 34 by the conditions

$$(6.12) \quad \alpha > 1, \quad \varepsilon(x) \leq \frac{K}{x^\delta} \quad (x \geq x_0, \quad x_0 \geq e), \quad \delta > 1, \quad K > 0.$$

Denote $\sqrt[\delta]{x}$ by x^ε and put in (5.6) $y = x^\varepsilon$. As by (6.12) $x \varepsilon(x^\varepsilon) \leq K$, it follows

$$(6.13) \quad \Lambda(x) \leq \alpha \Lambda(x^\varepsilon) + \beta K \quad (x \geq x_0).$$

Replacing here x by x^{ε^v} and multiplying by α^v we obtain

$$\alpha^v \Lambda(x^{\varepsilon^v}) \leq \alpha^{v+1} \Lambda(x^{\varepsilon^{v+1}}) + \beta K \alpha^v \quad (0 \leq v \leq n-1).$$

Adding over $v = 0, 1, \dots, n-1$ it follows

$$\Lambda(x) \leq \alpha^n \Lambda(x_0) + \beta K \sum_{v=0}^{n-1} \alpha^v \leq \alpha^n x_0 + \beta K \frac{\alpha^n - 1}{\alpha - 1} < \alpha^n \left(x_0 + \frac{\beta K}{\alpha - 1} \right).$$

as soon as $x^{\varepsilon^n} \leq x_0$, that is, as soon as

$$\lg x \leq \delta^n \lg x_0, \quad n \lg \delta \geq \lg \frac{\lg x}{\lg x_0}.$$

39. The last condition is satisfied as soon as $\text{nl}g\delta > \text{lg} \lg x$,

$$n > \frac{\text{lg} \lg x}{\text{lg} \delta} \quad , \quad n = \left[\frac{\text{lg} \lg x}{\text{lg} \delta} \right] + 1 \quad .$$

For this value of n it follows

$$\text{lg} \alpha^n = n \text{lg} \alpha < \text{lg} \alpha \left(1 + \frac{\text{lg} \lg x}{\text{lg} \delta} \right) = \text{lg} \left[(\text{lg} x)^{\frac{\text{lg} \alpha}{\text{lg} \delta}} \right] ,$$

or putting

$$(6.14) \quad \mu_0 := \frac{\text{lg} \alpha}{\text{lg} \delta} \quad , \quad \delta^{\mu_0} = \alpha \quad ,$$

$$\alpha^n < \delta^{\mu_0 (\text{lg} x)^{\mu_0}} = \alpha (\text{lg} x)^{\mu_0} \quad ,$$

Therefore, finally,

$$(6.15) \quad A(x) \leq \left(x_0 + \frac{\beta x}{\alpha - 1} \right) \alpha (\text{lg} x)^{\mu_0} \quad , \quad \mu_0 := \frac{\text{lg} \alpha}{\text{lg} \delta} \quad .$$

§ 7. $A(x)$ in dependence on $\Psi(x)$.

40. Returning now to the functional inequality (5.6) derived under the conditions specified in sec.1 we have to use the value (1.9) of $\xi(x)$,

$$(7.1) \quad \xi(x) = \frac{1}{\gamma_0 \varphi\left(\frac{1}{\gamma_0 x}\right)} \quad .$$

Thence, solving with respect to φ and using the inverse function Ψ ,

$$(7.2) \quad \varphi\left(\frac{1}{\gamma_0 x}\right) = \frac{1}{\gamma_0 \varepsilon(x)} \quad , \quad \frac{1}{\gamma_0 x} = \psi\left(\frac{1}{\gamma_0 \varepsilon(x)}\right) \quad .$$

However the cases (6.7) and (6.12) can be discarded.
Indeed under the assumption (6.7) it follows from (7.2)

$$\varphi\left(\frac{1}{\gamma_0 x}\right) = \frac{1}{\gamma_0 \varepsilon(x)} > \frac{x}{\gamma_0} \quad , \quad \psi\left(\frac{x}{\gamma_0}\right) > \frac{1}{\gamma_0 x}$$

so that finally $\psi(x) > \frac{1}{\gamma_0^2 x}$. But this is only possible for $m=1$,

$\gamma_0 = 1$ and in this case $\psi(x)$ is always $< 1/x$.

In the case of the condition (6.12) we obtain from (7.2)

$$\varphi\left(\frac{1}{\gamma_0 x}\right) > \frac{x^\delta}{\gamma_0 x} \quad , \quad \psi\left(\frac{x^\delta}{\gamma_0 x}\right) > \frac{1}{\gamma_0 x} \quad ,$$

and putting $y := x^\delta / (\gamma_0 x)$, $x = (\gamma_0 x y)^{1/\delta}$ it follows

$$\psi(y) > \frac{1}{\gamma_0 (\gamma_0 x y)^{1/\delta}}$$

which is impossible since $1/\delta < 1$.

We have therefore only to consider the case of sec. 34.

41. The assumptions (6.1), (6.2) and (6.3) in sec. 34
can be considerably simplified. Putting $g' > 1$,

$$(7.3) \quad 0 < \rho := \frac{1}{lg} \frac{g'}{g} < 1 \quad , \quad \varepsilon(x) := \left(\frac{\ell(x)}{x}\right)^\rho$$

the relation (6.2) becomes

$$\varepsilon(x) \leq g^f \varepsilon(gx), \left(\frac{\ell(x)}{x}\right)^f \leq g^f \left(\frac{\ell(gx)}{gx}\right)^f = \left(\frac{\ell(gx)}{x}\right)^f,$$

$$(7.4) \quad \ell(x) \leq \ell(gx) \quad (x \geq x_0).$$

The relations (6.1) and (6.3) become now

$$(7.5) \quad g > 1, 0 < f < 1, g^{1-f} > \alpha, L > 0, \varepsilon(x) > L \quad (1 \leq x \leq x_0).$$

The inequality (7.4) is in any case satisfied if $\ell(x)$ is assumed as non decreasing. In this particular case (6.2) holds for any sufficiently large $g > 1$ and, for a fixed f , (6.2) holds for all $g' = g^f$ from a $g > 1$ on. From now on we restrict ourselves to the case (7.3) with a constant, $0 < f < 1$.

The simplest case is of course $\ell(x) = c = \text{constans}$,

$$(7.6) \quad \varepsilon(x) \leq \left(\frac{c}{x}\right)^f, \quad 0 < f < 1, \quad r := \frac{1}{f}.$$

By (7.2) it follows

$$(7.7) \quad \varphi\left(\frac{1}{y_0 x}\right) \geq \frac{x^f}{y_0 c^f}, \quad \varphi(x) \geq \frac{y_0}{(c y^f x)^f}.$$

For the inverse function of $y = \varphi(x)$ it follows now

$$(7.8) \quad \Psi(y) \geq \left(\frac{y_0}{y}\right)^r \frac{1}{y_0^c}.$$

Inversely from (7.8) follows (7.7). From (6.5) we obtain now

$$(7.9) \quad \Lambda(x) = O(x^{1-f}).$$

42. The formula (7.9) holds in particular if the α_μ in (1.1) are all algebraic. To prove this denote by α some primitive element of the field $R(\alpha_1, \dots, \alpha_m)$ so that

$$\alpha_\mu = h_\mu(\alpha) \quad (\mu = 1, \dots, m)$$

where the h_μ are polynomials with rational coefficients. Then, denoting by u_0, u_1, \dots, u_m independent indeterminates, put

$$P(x) = \sum_{\mu=1}^m u_{\mu} h_{\mu}(x) + u_0.$$

Let be $n+1$ be the degree of α with respect to R . Denoting by $\alpha^{(v)}$ ($v = 0, 1, \dots, n$), $\alpha^{(0)} = \alpha$, the complete set of the conjugates of α , form the expression

$$T(u_0, \dots, u_m) = \prod_{v=0}^n P(\alpha^{(v)})$$

which is a polynomial with rational coefficients with common denominator N .

If we put for the u_{μ} rational integers g_{μ} with $\gamma := \max_{\mu} |g_{\mu}|$ we have for a fixed natural N :

$$N T(g_0, \dots, g_m) = G \neq 0,$$

with a rational integer G , so that $|N T(g_0, \dots, g_m)| \geq 1$.

On the other hand

$$T^*(g_0, \dots, g_m) := T(g_0, \dots, g_m) / \left[\sum_{\mu=1}^m g_{\mu} h_{\mu}(\alpha) + g_0 \right]$$

is of dimension n and therefore $T^*(g_0, \dots, g_m) = O(\gamma^n)$. It follows

$$\left| \sum_{\mu=1}^m g_{\mu} h_{\mu}(\alpha) + g_0 \right| > \frac{C}{\gamma^n}, \quad C > 0$$

with a constant C . We obtain from (1.2)

$$\Psi(\gamma) > \frac{C}{\gamma^n}$$

which is the relation (7.8) with $r = n$ and thence (7.9) with $\beta = 1/n$.

43. We can assume now $\ell(x)$ as strictly monotonically increasing. The essential difficulty in applying (6.5) consists in the necessity to obtain sufficiently good approximation of the inverse functions $\varphi(x)$ and $\psi(x)$. To do this we use the

Lemma 5. Assume $\ell(x)$ for an $x > x_0 > 1$ a positive strictly monotonically increasing function of x such that $x/\ell(x)$ also strictly monotonically

increases. Let $0 < \rho < 1$ and put $r := 1/\rho$. Then necessary and sufficient for the inequality

$$(7.10) \quad \xi(x) = \frac{1}{\ell_0 \varphi(1/\ell_0 x)} \leq c \left(\frac{\ell(x)}{x} \right)^\rho \quad (x \geq x_0)$$

is that $\psi(x)$ satisfies the inequality

$$(7.11) \quad \psi(y) \geq \frac{1}{\ell_0 (Dy)^{1/\rho} k((Dy)^{1/\rho})} \quad , \quad D := \ell_0^C \quad (y \geq y_0)$$

for a convenient constant $y_0 > 0$, where with

$$(7.12) \quad z := x / \ell(x) \quad , \quad z \geq \frac{x_0}{\ell(x_0)}$$

$k(z)$ is defined by

$$(7.13) \quad k(z) := \ell(x) \quad , \quad x = zk(z) \quad .$$

44. Proof. Using $\xi(x)$ from (2.10) it follows from (7.10)

$$(7.14) \quad \varphi\left(\frac{1}{\ell_0 x}\right) \geq \left(\frac{x}{\ell(x)}\right)^\rho / D =: y \quad .$$

Since $z = x/\ell(x)$ is strictly monotonically increasing, the same holds for $k(z)$ defined by (7.13) and it follows from (7.13) and (7.14) that

$$(7.15) \quad z = (Dy)^{1/\rho} \quad , \quad z \geq \frac{x_0}{\ell(x_0)} \quad , \quad y \geq \left(\frac{x_0}{\ell(x_0)}\right)^\rho / D =: y_0 \quad .$$

Applying to both sides of (7.14) the function ψ we obtain

$$\psi(y) \geq \frac{1}{\ell_0^C x}$$

and since by (7.13) and (7.15) $x = (Dy)^{1/\rho} k((Dy)^{1/\rho})$, (7.11) follows.

45. On the other hand, assuming (7.11) for $y \geq y_0$ with $(Dy_0)^{1/\rho} k((Dy_0)^{1/\rho}) > 1$ and defining z by (7.15) we can rewrite (7.11) as

$$(7.16) \quad \psi(y) \geq \frac{1}{\ell_0^{Ck(z)}} \quad , \quad z \geq (Dy_0)^{1/\rho} =: z_0 \quad .$$

Put then in (7.16)

$$(7.17) \quad x := z k(z), \quad x \geq z_0 k(z_0) =: x_0 > 1$$

and apply on both sides of (7.16) the function φ . We obtain

$$(7.18) \quad y \leq \varphi\left(\frac{1}{x_0 x}\right).$$

Defining now $\ell(x)$ by (7.13) we obtain from (7.12) and (7.15)

$$y = z^{\ell/D} = \left(\frac{x}{\ell(x)}\right)^{\ell/D}.$$

(7.18) becomes now

$$\varphi\left(\frac{L}{x_0 x}\right) \geq \left(\frac{x}{\ell(x)}\right)^{\ell/D}$$

and the formula (7.10) follows.

46. Applying the lemma 5 and starting from an inequality of the type of (7.11), it is important to find convenient functions $k(y)$. The following lemma allows this in a greater number of cases.

Lemma 6. Assume for $x \geq x_0 > e^e$, with $x \rightarrow \infty$:

$$(7.19) \quad e \leq k(x) \uparrow \infty, \quad (x \uparrow \infty)$$

$$(7.20) \quad x k'(x) = o(k(x)/\lg k(x)),$$

and define $Z(x)$ by:

$$(7.21) \quad z k(z) = x, \quad z = Z(x) \uparrow \infty.$$

Then for an arbitrary small $\varepsilon > 0$ with $x \rightarrow \infty$:

$$(7.22) \quad \frac{x}{k\left(\frac{x}{k(x)}\right)} > Z(x) > \frac{x}{k(x)},$$

$$(7.23) \quad k(x) = O(e^{\varepsilon \sqrt{\lg x}}),$$

$$(7.24) \quad Z(x) = \frac{x}{k(x)}(1 + o(1)).$$

47. Proof. From $k(x_0) > e$ it follows by (7.21) $Z(x_0) < x_0$, $Z(x) < x$.

$$k(z) < k(x) \quad (x \geq x_0)$$

and from

$$(7.25) \quad z = \frac{x}{k(z)}.$$

We obtain

$$(7.26) \quad Z(x) > \frac{x}{k(x)}, \quad k(z) > k\left(\frac{x}{k(x)}\right).$$

From (7.25) and (7.26) we obtain further

$$Z(x) < \frac{x}{k\left(\frac{x}{k(x)}\right)}$$

and (7.22) is proved.

48. By (7.20) we obtain, for an $\varepsilon > 0$,

$$x((\lg k(x))^2)' < \varepsilon^2 \quad (x > x_1),$$

$$((\lg k(x))^2)' < \frac{\varepsilon^2}{x} \quad (x > x_1),$$

$$(\lg k(x))^2 < \varepsilon^2 \lg x + (\lg k(x_1))^2 - \varepsilon^2 \lg x_1,$$

$$\lg k(x) < \sqrt{\varepsilon^2 \lg x + c} < \varepsilon \sqrt{\lg x} + \sqrt{c}$$

for a constant c , and (7.23) follows.

49. Finally using (7.20) we obtain

$$k(x) - k\left(\frac{x}{k(x)}\right) = \int_{x/k(x)}^x k'(y) dy = o\left(\int_{x/k(x)}^x \frac{k(y)}{\lg k(y)} \frac{dy}{y}\right).$$

But obviously, in virtue of

$$(\lg k(x))^2 \left(\frac{k(x)}{\lg k(x)} \right)' = k'(x) (\lg x - 1) \geq 0,$$

we can take in the last integral the factor $\frac{k(x)}{\lg k(x)}$ out of the integral and obtain

$$k(x) - k\left(\frac{x}{k(x)}\right) = o\left(\frac{k(x)}{\lg k(x)}\right) \int_{x/k(x)}^x \frac{dy}{y} = o\left(\frac{k(x)}{\lg k(x)} \lg k(x)\right) = o(k(x)).$$

It follows

$$k(x)/k\left(\frac{x}{k(x)}\right) \rightarrow 1 + o(1)$$

and (7.24) follows from (7.22). Lemma 6 is proved.

50. We can now formulate in a particularly simple and important case
Theoreme 2. Assume $k(x)$ a constant or strictly increasing function
satisfying the conditions (7.19) and (7.20). Assume (7.11) for a convenient
 $C > 1$ and a φ with $0 < \varphi < 1$. Then

$$(7.28) \quad \mathcal{E}(x) = o\left(\left(\frac{k(x)}{x}\right)^\varphi\right),$$

$$(7.29) \quad \Lambda(x) = o(x^{1-\varphi} k(x)^\varphi).$$

Proof. Defining z by (7.21) it follows from $\ell(x) := k(z)$ and (7.24):

$$\frac{\ell(x)}{x} = \frac{k(z)}{z k(z)} = \frac{1}{z} = \frac{k(x)}{x} (1 + o(1)),$$

$$\ell(x) = k(x)(1 + o(1))$$

and therefore (7.28) and (7.29).

51. Consider, for instance, the monotonically increasing expressions of the type

$$k(x) := c \lg_1^{\alpha_1} x \lg_2^{\alpha_2} x \dots \lg_n^{\alpha_n} x \quad (x > x_0),$$

where generally the ν -times iterated logarithm of x is denoted by $\lg_\nu x$ and the first non vanishing term in the sequence $\alpha_1, \alpha_2, \dots, \alpha_n$ is positive.

Then we have for the logarithmic derivative of $k(x)$:

$$\frac{k'(x)}{k(x)} = \sum_{v=1}^n \frac{\alpha_v}{x \lg_1 x \dots \lg_v x} = O\left(\frac{1}{x \lg x}\right).$$

Since $\lg k(x) = O(\lg_2 x)$ it follows

$$x k'(x) \lg k(x) / k(x) = O\left(\frac{\lg_2 x}{\lg x}\right) = o(1)$$

and the conditions of lemma 6 are satisfied. It follows

$$(7.30) \quad \Lambda(x) = O\left(\frac{x^{1-\delta}}{k(x)}\right).$$

Bibliography

- [1] Cassels, J.W.S. , An Introduction to Diophantine Approximation, Cambridge Tracts, Cambridge University Press, 1951.
- [2] Hartman, S. , Une Généralisation d'un Théorème de M. Ostrowski sur la répartition des nombres mod 1. Ann.Soc.Pol.Math.vol.22, pp.169-172,(1949).
- [1] Hecke, E. , Über analytische Funktionen und die Verteilung von Zahlen mod eins, Hamb.Math.Abh.,1 (1921), pp. 54-76, in particular pp. 72-76. Reprinted in Hecke's Math. Werke, Göttingen 1959, pp.313-335, in particular § 6, pp. 331-335.
- [1] Khintchine, A. , The quantitative conception of Kronecker's approximation theory, Isv.Acad. USSR 12 (1943), pp. 113-122.
- [1] Ostrowski, A. , Zur Theorie der linearen Diophantischen Approximationen, Jahresbericht d.D.M.V.,XXXIX, 1930, pp.34-46.
- [2] Ostrowski, A. , On rational approximations to an irrational number, BEM 50.
- [1] Weyl, H. , Über die Gibbsche Erscheinung und verwandte Konvergenzphänomene, Rend.Circ.Mat.Palermo, 30, 1910, pp. 377-407, in particular p. 406. Reprinted in Weyl's collected papers, vol.1, Springer 1968, pp.352-353.

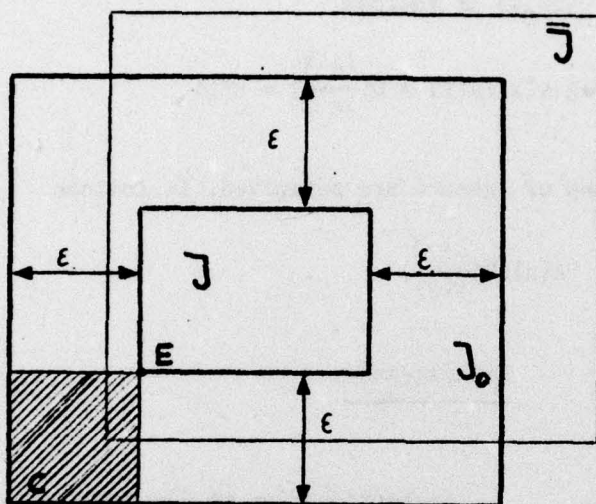


Fig. 1

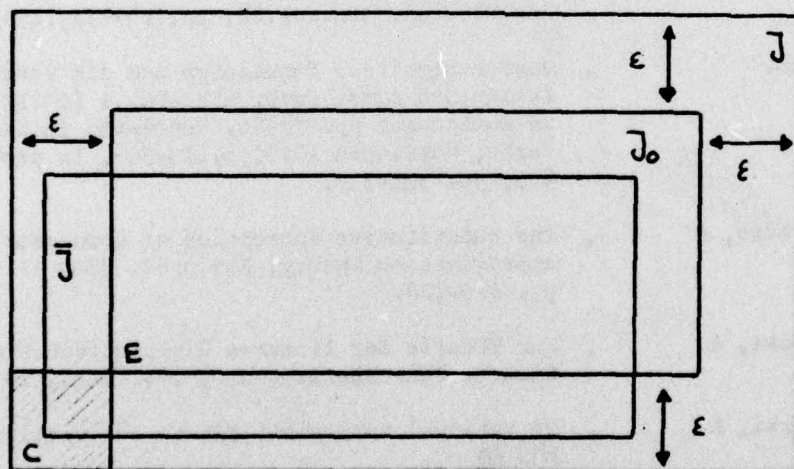


Fig. 2